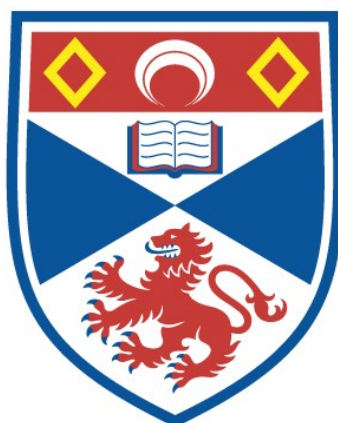


BIOSYNTHESIS AND METABOLISM  
OF SELECTED ORGANOFLUORINE COMPOUNDS

Andrea Rodil Garcia

A Thesis Submitted for the Degree of PhD  
at the  
University of St Andrews



2019

Full metadata for this item is available in  
St Andrews Research Repository  
at:

<http://research-repository.st-andrews.ac.uk/>

Please use this identifier to cite or link to this item:

<http://hdl.handle.net/10023/18485>

This item is protected by original copyright

# Biosynthesis and Metabolism of Selected Organofluorine Compounds

Andrea Rodil Garcia



University of  
St Andrews

This thesis is submitted in partial fulfilment for the degree of  
Doctor of Philosophy (PhD)  
at the University of St Andrews

February 2019



## Candidate's declaration

I, Andrea Rodil Garcia, do hereby certify that this thesis, submitted for the degree of PhD, which is approximately 58,000 words in length, has been written by me, and that it is the record of work carried out by me, or principally by myself in collaboration with others as acknowledged, and that it has not been submitted in any previous application for any degree.

I was admitted as a research student at the University of St Andrews in September 2014.

I received funding from an organisation or institution and have acknowledged the funder(s) in the full text of my thesis.

Date \_\_\_\_\_ Signature of candidate \_\_\_\_\_

## Supervisor's declaration

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of PhD in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree.

Date \_\_\_\_\_ Signature of supervisor \_\_\_\_\_

## Permission for publication

In submitting this thesis to the University of St Andrews we understand that we are giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. We also understand, unless exempt by an award of an embargo as requested below, that the title and the abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker, that this thesis will be electronically accessible for personal or research use and that the library has the right to migrate this thesis into new electronic forms as required to ensure continued access to the thesis.

I, Andrea Rodil Garcia, confirm that my thesis does not contain any third-party material that requires copyright clearance.

The following is an agreed request by candidate and supervisor regarding the publication of this thesis:

## Printed copy

No embargo on print copy.

## Electronic copy

No embargo on electronic copy.

Date \_\_\_\_\_ Signature of candidate \_\_\_\_\_

Date \_\_\_\_\_ Signature of supervisor \_\_\_\_\_

## Underpinning Research Data or Digital Outputs

## Candidate's declaration

I, Andrea Rodil Garcia, hereby certify that no requirements to deposit original research data or digital outputs apply to this thesis and that, where appropriate, secondary data used have been referenced in the full text of my thesis.

Date \_\_\_\_\_ Signature of candidate \_\_\_\_\_

*A Clara*

*Some people are worth melting for  
And you will always be one of them,  
I would melt a thousand times for you if I could.*

*Sleep tight, our little princess,  
thank you for those wonderful eight years you gave us.*

*You will always be in my heart.*

*This thesis is for you, with all my love*

## Acknowledgements

First, and most importantly, I would like to thank Professor David O'Hagan for the opportunity he gave me when he first accepted me as a PhD candidate in his group. Not only I have had the chance to come to the beautiful St. Andrews – and to discover Scotland – but it has also been an absolute honour to work under the guidance of such a passionate chemist. Thank you, David!

This project was funded – and therefore possible – by the Marie Curie 7<sup>th</sup> Framework, and the European Union, as part of the Fluor21 Network. As part of the Fluorine21 Network, I have been lucky to meet a very colourful array of scientists. Particularly grateful I am to Dr. Cormac Murphy, Dr. Marta Saccomanno and Dr. William Palmer-Brown, for everything they taught me about *C. elegans*, but also for the very especial time spent in Dublin with them! In general, many thanks to Etienne, Tom, Rita, Maté, Diana, Dave and Jorg, and obviously all the PIs, for the fantastic meetings together! Within the network, my most loving thanks to my fluorinated elder sister and brother, Anna and Javi. I love you so much!

Most of the results in this work were fathomed thanks to the analytical departments of the School of Chemistry. Thank you to Mrs. Melanja Smith, Dr. Tomas Lebl and Dr. Siobhan Smith for the NMRs; to Dr. Catherine Botting and Dr. Sally Shirran for protein mass spec; and to Mrs. Caroline Horsburgh for all the high-resolution mass spectrometry. My gratitude goes as well to Prof. Alex Slawin and Dr. David Cordes for the X-rays.

Great part of this research would not have been possible without the help of many excellent scientists. The O'Hagan group (past and present) has made the task of finishing a PhD a little easier, so thanks to all of you guys! Thank you for every single contribution you have made to this work. My especial gratitude goes to Dr. Stephen Thompson, for his infinite patience with me, and his (seemingly) never-ending chemical wisdom. It was such a pleasure to have you as a labmate! Dr. Qingzhi Zhang, Qing, thank you so much for your support through the hardships, and your help with the chemistry; basically, thank you for being my resident mum in the lab! To Dr.

Nouchali Bandaranayaka, many, many thanks for all the times you have been there for me, as a colleague, but mostly as a friend, I will never forget it.

This experience would not have been the same without the lunches in the ‘Spanish Table’. Thanks to all of you, Juanma, Rodri, Susana, Diego, Jesús, Iva and Angela!

I would most certainly not be here, writing this page, if I hadn’t counted with the support of the St Andrews Shinty Team. Thank you guys! Keep being amazing! Within the team, especial mention should go to Hestia, Fauna and Mopsee, for all the amazing times – and brunches – together. Thank you for being such wonderful friends. Mopseeta, wherever you go next, please keep telling me all your fascinating stories. Couldn’t have done this without you.

Double the gratitude, admiration, respect and love go to Dobbee – or should I say Felix? –, that unlikely link between Shinty and Church. Not only were you always there to offer your friendly hand when most needed, but you also guided me back to where I belong. I will be grateful forever.

To all the family at St Andrews Free Church, for being a constant source of support, wisdom, guidance and friendship (quite often while feeding me amazing dishes!). Abigail, Bethan, Jody, Kirsty, Haley, Rachel, Sarah K., Sarah C., Anneleen, Shaunagh, Emily, Annabel and Alex, I hope our paths keep crossing! To all the Church family that I am not mentioning here: I am so grateful to have you all, I have been indeed blessed with you. Thank you for the amazing community we have in St Andrews!

In this last year of my thesis, I have had the good fortune of starting to work for CAPOD, where I met really great colleagues – and bosses – like Cat and Rebecca, but also the other half of the ‘Madras Team’, the incredible Francesca. Thank you so much for being such an amazing friend, it was great fun to work with you!

Back in Spain I’ve counted with the support of the Guitián-Pérez-Peña group, who introduced me to the wonders of chemistry. And I will never look back! Thank you to Dolores for being such an amazing first supervisor, and thank you so much to the rest of the group for making it a real family. Los ‘Guitianes’ (Jenni, Manu y Sabe) will always have a special place in my heart, along with Sara, Manu Vilas and Diego Lojo.

Mil gracias a Marta, por seguir siendo mi mejor amiga a pesar de la distancia y el tiempo, no ha sido fácil. Y gracias a María y Ana, mis porquis, por ayudarme a sobrevivir Química, y por las visitas a St Andrews. Seguid estando a mi lado chicas.

Finalmente, gracias a mi familia, por cuidarme y entenderme, y por darme la libertad de poder seguir mi propio camino, por muy lejos de casa que me lleve. Mamá y papá, ¡os quiero tanto! Este doctorado es para vosotros. Buelita, tu cara cada vez que me ves por sorpresa en la puerta es el mejor regalo, gracias. Tata, gracias por hacer de prima mayor y mimarme tanto. A todos mis tíos y primos, Tere, Lucía, Amparo, Artemio, Carlos, Carlitos y Piti, a vosotros también gracias. Especialmente a Nina y Alberto, por dejarme ser su sobrina consentida. Nory, Alber, Rodri, Rosa, Ana y Fernando, y todos los “peques” (Santi, Fer, Rocío, Ani, Clara, Sofi y Rodri): gracias por estar siempre ahí. Os quiero.

## Abstract

The St Andrews group has recently developed a range of novel fluorinated motifs, which include all-*cis* fluorocyclohexanes, difluoro oxy- and thio- ethers and  $\alpha, \beta, \beta$ -trifluorocyclopropanes. Fluorinated motifs are often sought after in medicinal chemistry, and it became one of the aims of this thesis to explore these new motifs in order to assess their validity for drug discovery programmes.

The metabolic stability of each motif was probed in incubations with *Cunninghamella elegans*. This fungus is well known for its ability to biotransform xenobiotics mimicking the mammalian P<sub>450</sub> profile, and thus, providing a useful model for metabolic vulnerability. This showed that the all-*cis* cyclohexyl derivatives were increasingly stable to the action of P<sub>450</sub>'s with increasing fluorine substitution. The  $\alpha, \alpha$ -difluoro thioethers showed oxidation to sulfoxides and sulfones, and the difluorooxy analogue showed a rapid biodegradation into the corresponding acetate ester. The studied  $\alpha, \beta, \beta$ -trifluorocyclopropanes proved challenging to isolate and identify, but it could be established that the fluorinated motif remained intact after the incubations. The enantioselectivity with which *C. elegans* transforms the products was explored by chiral HPLC, which allowed an estimation of the enantiomeric resolution of racemates, and also the enantioselectivity of sulfoxide formation. In both cases selectivities were relatively low. Additionally, the lipophilicity of the fluorinated motifs was calculated by reverse-phase HPLC. A general trend was that increasing fluorination led to increased polarity.

A second aim was to explore the fluorometabolite biosynthesis in *Streptomyces* sp. MA37, isolated in Ghana (2014), which contains an active fluorinase. The metabolic pathway in *Streptomyces* sp. MA37 branches at 5-fluorodeoxyribose phosphate (5-FDRP), allowing the formation of 5-fluoro-2,3,4-trihydroxypentanoic acid (5-FHPA) and other unknown metabolites. Each of the steps leading to 5-FHPA was studied, and the biotransformation was reconstituted *in vitro*.

The production of fluorometabolites by *Streptomyces* sp. MA37 was investigated by carrying out cell-free and whole cell feeding experiments with [5',5'-<sup>2</sup>H<sub>2</sub>]-5'-FDA. Incorporation into the fluorometabolites from [5',5'-<sup>2</sup>H<sub>2</sub>]-5'-FDA was successful.

## Abbreviations

[ $\alpha$ ] <sub>D</sub>	Specific rotation
Ac	Acetyl group
AcCN	Acetonitrile
ADA	Adenosine deaminase
ADME	Absorption, Distribution, Metabolism, Excretion
AMP	Adenosine monophosphate
Asp	Aspartate
Atm.	Atmosphere
ATP	Adenosine triphosphate
BAIB	Bis(acetoxy)iodobenzene
Bu	Butyl group
°C	Celsius degrees
c.c.	Column chromatography
CFE	Cell-free extract
CH <sub>3</sub> CN	Acetonitrile
CIDA	Chlorodeoxyadenosine
CNS	Central Nervous System
CoA	Coenzyme A
CPR	Cytochrome P <sub>450</sub> Reductase
CYP	Cytochroma P <sub>450</sub>
DMP	2,2-Dimethoxypropane
D	Debye
Da	Dalton

$\delta$	Chemical shift (expressed in ppms)
$\Delta$	Heating, heat
d	Doublet
dd	Doublet of doublets
DAST	Diethylaminosulfur trifluoride
DBU	1,8-Diazabicyclo[5.4.0]undec-7-ene
DDT	Dichlorodiphenyl trichloroethane
DMF	Dimethylformamide
duf	Domain of unknown function
E	Electrophile
e.e.	Enantiomeric excess
eq. or equiv.	Equivalent
ESI	Electrospray
Et	Ethyl group
Et <sub>3</sub> N	Triethylamine
FAc	Fluoroacetate
FAD	Flavin adenosine dinucleotide
FAld	Fluoroaldehyde
5'-FDA	5'-Fluoro-5-deoxyadenosine
FDG	Fluorodeoxyglucose
5'-FDI	5'-Fluoro-5-deoxyinosine
5-FDR	5-Fluoro-5-deoxyribose
5-FDRP	5-Fluoro-5-deoxyribose phosphate
FDRibulP	Fluorodeoxyribulose phosphate
5-FHPA	5-Fluoro-2,3,4-trihydroxypentanoic acid
fid	Free induction decay
Fig.	Figure



<i>flA</i>	Fluorinase
<i>flK</i>	Fluoroacetyl-CoA thioesterase
FMN	Flavin mononucleotide
4-FT	4-Fluorothreonine
g	Gram(s)
h	Hour(s)
His	Histidine
HPLC	High-performance liquid chromatography
Hz	Hertz(s)
<i>I</i>	<i>Iso</i>
IPTG	<i>Iso</i> -propyl- $\beta$ -D-1-thiogalactopyranoside
IR	Infrared
<i>J</i>	Coupling constant (expressed in Hz)
l	Liquid
L	Litre
LB	Lysogeny broth
LC-MS	Liquid chromatography-mass spectrometry
LDH	Lactate dehydrogenase
<i>L</i> -Met	<i>L</i> -methionine
m	Mass, multiplet
<i>m</i>	<i>Meta</i>
M	Molar
MALDI	Matrix-assisted laser desorption ionisation
<i>m</i> CPBA	<i>Meta</i> -chloroperbenzoic acid
Me	Methyl group
MeOH	Methanol

mg	Milligram
MHz	Megahertz
min.	Minute
μL	Microlitre
mL	Millilitres
mmol	Minimol
MOPS	3-( <i>N</i> -morpholino)propanesulfonic acid
m.p.	Melting point
<i>m/z</i>	Mass to charge ratio
NAD <sup>+</sup>	Nicotinamide adenine dinucleotide
NADH	Nicotinamide adenine dinucleotide (reduced)
NADPH	Nicotinamide adenine dinucleotide phosphate (reduced)
NBS	N-Bromosuccinimide
NH	Nucleoside hydrolase
Ni-NTA	Nickel-Nitrilotriacetic acid (agarose resin)
NMR	Nuclear magnetic resonance
NS	Number of scans
Nu	Nucleophile
<i>o</i>	<i>Ortho</i>
OD	Optical density
O1P	Frequency offset point
orf	Open Reading Frame
<i>p</i> -TsOH	<i>Para</i> -toluenesulfonic acid
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PET	Positron emission tomography
PLP	Pyridoxal phosphate

PNP	Purine nucleoside phosphorylase
PP <sub>i</sub>	Pyrophosphate
PTFE	Polytetrafluoroethylene
ppm	Parts-per-million
q	Quartet or quintet
rpm	Revolutions per minute
r.t.	Room temperature
R <sub>t</sub>	Retention time
s	Singlet
SAM	S-Adenosyl methionine
sat.	Saturated
SDS-PAGE	Sodium dodecylsulfate polyacrylamide gel electrophoresis
t	Triplet
t <sub>1/2</sub>	Half life
TBAF	Tetrabutylammonium fluoride
TCA	Tricarboxylic acids (cycle)
TEMPO	(2,2,6,6-tetramethylpiperidin-1-yl)oxidanyl
TEV	Tobacco Etch Virus
Tf	Triflate (Trifluoromethanesulfonyl)
TFA	Trifluoroacetic acid
TFE	2,2,2-Trifluoroethanol
TFE	Tetrafluoroethylene (only mentioned in Chapter 1)
THF	Tetrahydrofuran
TLC	Thin layer chromatography
Tris·HCl	2-amino-2-hydroxymethylpropane-1,3-diol – hydrochloric acid
Ts	Tosyl (Toluenesulfonyl)

TsF	Tosyl fluoride ( <i>para</i> -toluenesulfonic fluoride)
<i>p</i> -TsOH	<i>para</i> -Toluenesulfonic acid
T <sub>r</sub> or T <sub>ret</sub>	Retention time
SNRI	Serotonin and Norepinephrine Re-uptake Inhibitor
UV	Ultraviolet

# Table of Contents

<b>Chapter 1. Introduction.....</b>	<b>6</b>
<b>1.1 A brief history of fluorine.....</b>	<b>6</b>
1.1.1 Discovery, early uses and applications.....	6
1.1.2 Fluorine in agrochemicals.....	8
1.1.3 Fluorine in materials chemistry .....	11
1.1.4 Fluorine in PET .....	13
1.1.5 Fluorine in the pharmaceuticals and medicinal chemistry .....	14
<b>1.2 Metabolism of Drugs .....</b>	<b>21</b>
<b>1.3 <i>Cunninghamella elegans</i> as a model for mammalian drug metabolism.....</b>	<b>23</b>
1.3.1 The fungus <i>Cunninghamella elegans</i> .....	23
1.3.2 Cytochrome P <sub>450</sub> and metabolism .....	24
1.3.3 Biotransformation of xenobiotics .....	26
<b>1.4 Natural Products. Relevance and biosynthesis.....</b>	<b>29</b>
1.4.1 Relevant natural products. Discovery, roles and applications .....	29
1.4.2 Halogenated natural products and enzymatic halogenation .....	30
1.4.3 Fluorinated natural products .....	38
<b>1.5 The fluorinase .....</b>	<b>47</b>
1.5.1 Discovery and crystal structure of the first fluorinating enzyme .....	47
1.5.3 <i>Streptomyces</i> sp. MA37 .....	49
1.5.4 Other organisms containing a fluorinase .....	52
<b>1.6 References.....</b>	<b>53</b>
 <b>Chapter 2. Metabolism and lipophilicity profile of a range of novel fluorinated motifs .....</b>	 <b>65</b>
<b>2.1 Derivatives of all-<i>cis</i> phenylfluorocyclohexane.....</b>	<b>66</b>
2.1.1 Introduction to the syntheses and potential applications of all- <i>cis</i> (fluorocyclohexyl)benzenes .....	66
2.1.2 Aims and Objectives.....	74
2.1.3 Estimation of lipophilicities by the use of <sup>19</sup> F NMR spectroscopy .....	74
2.1.4 Determination of Log <i>P</i> values by reverse-phase HPLC.....	81
2.1.5 Metabolic fate in <i>Cunninghamella elegans</i> .....	87
2.1.6 Enantiomeric excess analysis of all- <i>cis</i> (fluorocyclohexyl)benzene derivatives .....	92

2.1.7 Conclusions on the work with all- <i>cis</i> (fluorocyclohexyl) benzenes .....	97
<b>2.2 <math>\alpha,\alpha</math>-Difluoroethyl thioethers and <math>\alpha,\alpha</math>-difluoro oxyethers .....</b>	<b>98</b>
2.2.1 Introduction to the syntheses and potential applications of $\alpha,\alpha$ -difluoroethyl thio- and oxy- ethers.....	98
2.2.2 Aims and Objectives.....	100
2.2.3 Estimation of lipophilicities by the use of a reverse-phase HPLC method .....	100
2.2.4 Metabolic fate of the ethers in <i>Cunninghamella elegans</i> .....	107
2.2.5 Enantiomeric excess analysis .....	112
2.2.6 Conclusions on $\alpha,\alpha$ -difluoroethyl thio- and oxy- ethers.....	115
<b>2.3 <math>\alpha,\beta,\beta</math>-(Trifluorocyclopropyl) benzene derivatives.....</b>	<b>116</b>
2.3.1 Introduction to the syntheses and potential applications of fluorocyclopropanes.....	116
2.3.2 Aims and Objectives.....	118
2.3.3 Estimation of lipophilicities by the use of a reverse-phase HPLC .....	119
2.3.4 Metabolic fate of partially fluorinated cyclopropanes in <i>Cunninghamella elegans</i> .....	121
2.3.5 Conclusions on $\alpha,\beta,\beta$ -(trifluorocyclopropyl) aryl derivatives.....	127
<b>2.4 General Conclusions on Chapter 2 .....</b>	<b>128</b>
<b>2.5 References .....</b>	<b>130</b>
<b>Chapter 3. <i>In vitro</i> reconstitution of (2<i>R</i>,3<i>S</i>,4<i>S</i>)-5-fluoro-2,3,4-trihydroxypentanoic acid biosynthesis .....</b>	<b>135</b>
3.1 Introduction to (2 <i>R</i> ,3 <i>S</i> ,4 <i>S</i> )-5-fluoro-2,3,4-trihydroxypentanoic acid (2) biosynthesis .....	135
3.2 Aims and Objectives.....	138
3.3 <i>In vitro</i> reconstitution of 5-FHPA biosynthesis.....	138
3.3.1 Transformation of fluoride to 5-FHPA (2) .....	138
3.3.2 Individual enzymatic analysis of the biotransformation from fluoride to 5-FHPA (2).....	143
3.4 Optimisation of the enzymatic route.....	150
3.4.1 Using a nucleoside hydrolase .....	150
3.4.2 Implementation of an NAD <sup>+</sup> recycling protocol .....	153
3.5 Scale-up of the 5-FHPA biotransformation .....	157
3.6 Conclusions .....	164
3.7 References .....	165

<b>Chapter 4. Exploring the fluorinase pathway of <i>Streptomyces</i> sp. MA37</b>	<b>167</b>
<b>4.1 Introduction</b>	<b>167</b>
<b>4.2 Aims and objectives</b>	<b>171</b>
<b>4.3 Fluorometabolite profile in <i>Streptomyces</i> sp. MA37</b>	<b>172</b>
<b>4.4 Cell-free experiments with <i>Streptomyces</i> sp. MA37</b>	<b>173</b>
4.4.1 Exploratory incubations of SAM with inorganic fluoride, and 5'-FDA.	
Assignment of signals	173
4.4.2 Effect of the addition of fluorinase and influence of 5'-FDA concentration	177
4.4.3 Addition of cofactors, supplements and enzymes into the CFE incubations	181
4.4.4 Incubation of 5-FDR. Reproduction of published experiments	184
<b>4.5 Incubation in cell-free experiments from <i>Streptomyces cattleya</i></b>	<b>184</b>
<b>4.6 Synthesis of the deuterium labelled intermediates for feeding experiments</b>	<b>185</b>
4.6.1 Synthesis of [5',5'- <sup>2</sup> H <sub>2</sub> ]-5'-fluoro-5'-deoxyadenosine (17a) and [5,5- <sup>2</sup> H <sub>2</sub> ]-5'-fluoro-5-deoxyribose (26a)	186
<b>4.7 Feeding experiments in whole cell cultures of <i>S. sp.</i> MA37</b>	<b>190</b>
4.7.1 Feeding of 5'-FDA (17)	190
4.7.2 Feeding of [5',5'- <sup>2</sup> H <sub>2</sub> ]-5'-FDA (17a)	191
<b>4.8 Conclusions</b>	<b>195</b>
<b>4.9 References</b>	<b>196</b>
<b>Chapter 5. Experimental</b>	<b>197</b>
<b>5.1 General Procedures and Methods</b>	<b>197</b>
5.1.1 Chemistry general procedures	197
5.1.2 Biology general procedures	199
<b>5.2 Synthetic preparation of compounds</b>	<b>201</b>
5.2.1 Racemic ((1R,2R,3R,6S)-1-bromo-2,3,6-trifluorocyclohexyl)benzene (91)	201
5.2.2 Racemic (1R,2S,3R,6S)-2,3,6-trifluoro-1-phenylcyclohexan-1-ol (79)	202
5.2.3 Racemic methyl 4-((1S,2S,3R,6S)-2,3,6-trifluorocyclohexyl)benzoate (93)	203
5.2.4 ( <i>E</i> )-((1,2-dichlorovinyl)oxy)benzene (108)	204
5.2.5 (Ethynyloxy)benzene (109)	205
5.2.6 (1,1-Difluoroethoxy)benzene (101)	206
5.2.7 1-((1,1-Difluoroethyl)sulfonyl)-4-methoxybenzene (115)	207
5.2.8 Racemic 1-((1,1-difluoroethyl)sulfinyl)-4-methoxybenzene (112)	208
5.2.9 Racemic 2-((1,1-difluoroethyl)sulfinyl)naphthalene (116)	209

5.2.10 2,2-Difluoro-1-(4-methoxyphenyl)cyclopropan-1-ol (152).....	210
5.2.11 2',3'- <i>O</i> -Isopropylideneadenosine (155).....	211
5.2.12 2',3'- <i>O</i> -Isopropylidene-5'-fluoro-5'-deoxyadenosine (156).....	212
5.2.13 5'-Fluoro-5'-deoxyadenosine (5'-FDA, 17) .....	213
5.2.14 2,3- <i>O</i> -Isopropylidene-1-methoxy ribose (159).....	214
5.2.15 2,3- <i>O</i> -Isopropylidene-1-methoxy-5-tosyl ribose (160).....	215
5.2.16 5-Fluoro-5-deoxy-2,3- <i>O</i> -isopropylidene-1-methoxy ribose (161) .....	216
5.2.17 5-Fluoro-5-deoxoribose (5-FDR, 26).....	217
5.2.18 2',3'- <i>O</i> -Isopropylidene adenosine-5'-carboxylic acid (167).....	218
5.2.19 Adenosine-5'-carboxylic acid methyl ester (169).....	219
5.2.20 5',5'- <sup>2</sup> H <sub>2</sub> -Adenosine (154a) .....	220
5.2.21 [5',5'- <sup>2</sup> H <sub>2</sub> ]-2',3'- <i>O</i> -Isopropylideneadenosine (155a) .....	221
5.2.22 5'-Deoxy-5'-fluoro-5',5'- <sup>2</sup> H <sub>2</sub> -2',3'- <i>O</i> -isopropylideneadenosine (156a) .....	222
5.2.23 5'-Deoxy-5'-fluoro-5',5'- <sup>2</sup> H <sub>2</sub> -adenosine (17a).....	223
<b>5.3 Fluorometabolite production in <i>Cunninghamella elegans</i> .....</b>	<b>224</b>
5.3.1 General procedures for the works with <i>C. elegans</i> .....	224
5.3.2 Fluorometabolite production in <i>C. elegans</i> in all- <i>cis</i> fluorocyclohexyl benzene derivatives.....	225
5.3.3 Fluorometabolite production in <i>C. elegans</i> in $\alpha,\alpha$ -difluoroethyl thio- and oxy- ethers.....	232
5.3.4 Fluorometabolite production in <i>C. elegans</i> in $\alpha,\beta,\beta$ -(trifluorocyclopropyl) benzene derivatives .....	237
<b>5.4 Enantiomeric excess analysis .....</b>	<b>238</b>
5.4.1 Enantiomeric excess analysis in all- <i>cis</i> fluorocyclohexyl derivatives .....	238
5.4.2 Enantiomeric excess analysis for $\alpha,\alpha$ -difluoroethyl thioethers.....	241
<b>5.5 Estimation of lipophilicities .....</b>	<b>243</b>
5.5.1 Estimation by <sup>19</sup> F NMR .....	243
5.5.2 Estimation by reverse-phase HPLC.....	250
<b>5.6 Enzymatic biotransformations (Chapter 3).....</b>	<b>256</b>
5.6.1 Preparation of buffers.....	256
5.6.2 Media preparation.....	259
5.6.3 Over-expression and purification of nucleoside hydrolase (NH) <sup>6</sup> .....	259
5.6.4 Transformation, over-expression and purification of FdrC <sup>2</sup> .....	260
5.6.5 Individual enzymatic assays .....	263
5.6.6 Experiments via phytase.....	264



5.6.7 Experiments <i>via</i> NH .....	266
5.6.8 Biotransformations with LDH as recycling agent for NAD <sup>+</sup> .....	267
5.6.9 Biotransformations with FMN as recycling agent for NAD <sup>+</sup> .....	268
5.6.10 Scale-up of the biosynthesis of 5-FHPA (2) from 5'-FDA (17).....	270
<b>5.7 Fluorometabolite studies in <i>Streptomyces</i> sp. MA37 (Chapter 4) .....</b>	<b>271</b>
5.7.1 Media preparation .....	271
5.7.2 Preparation of cell-free extracts (CFE) from <i>S. sp.</i> MA37 .....	272
5.7.3 Preparation of a cell-free extract of <i>S. cattleya</i> .....	274
5.7.4 5'-FDA (17) enzymatic transformations (Section 4.4.1).....	275
5.7.5 Incubations in whole cell cultures of <i>S. sp.</i> MA37 .....	276
<b>5.8 References .....</b>	<b>278</b>
<b>Appendix – Crystallographic Data .....</b>	<b>282</b>
<b>Appendix – Publications .....</b>	<b>288</b>

# Chapter 1. Introduction

## 1.1 A brief history of fluorine

### 1.1.1 Discovery, early uses and applications

Fluorine was first isolated in 1886 by the French chemist Henri Moissan.<sup>1,2</sup> His achievement had been unsuccessfully attempted by many chemists before – including his mentor Frémy –,<sup>3</sup> and that granted Moissan a Nobel Prize in 1906.<sup>4</sup> However, an awareness of fluorine preceded these events.

The name ‘fluorine’ was first used in 1812 (almost a century before Moissan’s achievement) by another French scientist, André Ampère.<sup>5</sup> Ampère based his choice on the description of the use of fluorspar given by Georgius Agricola, in 1529. Agricola, a German scientist known in history as ‘the father of mineralogy’, noted that fluorspar could be used as a flux for molten metals in his most famous work, ‘*De re metallica*’.<sup>6</sup> Agricola described how the fluorinated mineral melted and flowed when heated in a fire, and when added to metals, their melt would become more fluid. Therefore, Ampère took the Latin word ‘*fluere*’, meaning ‘to flow’, to name this new element in its molecular form, F<sub>2</sub>. It is said that he later regretted his choice, in favour of another characteristic of fluorine: destruction. He found the name ‘*phthorine*’, from the Greek word for destructive ‘*phthoros*’, much more appropriate.<sup>7</sup> However, the name fluorine had been popularised already and other chemists, like Humphry Davy, were using it,<sup>8</sup> so it secured its place in history.

The element is found in its mineral forms, fluorite (or fluorspar), apatite and cryolite, as fluoride. Hence, Agricola might have been the first person to record the properties of fluorspar, but this fluorine-containing mineral had been used long before the Middle Ages.

Egyptians were probably the first civilisation to use calcium fluoride. Its uses were mainly ornamental, being shaped into beads and vessels.<sup>9</sup> Werner and Bimson (1963)<sup>10</sup> described fluorite as an ‘opacifying material’ used as an imitation for jade during the

times of the Tang dynasty (AD 618-907). In this context it was very specific to some regions of South-East Asia.<sup>11</sup>

The use of fluorspar increased considerably, and according to McLaughlin, fluorite coming from the desserts of Iran became the most popular crystal for Roman society.<sup>12</sup> Apparently, during the early period of the Empire, the finest Iranian fluorspar goblets could command a price higher than a city mansion.<sup>12</sup> One of the most famous examples of these goblets might be ‘The Barber Cup’, currently displayed in the British Museum in London. This one-handed fluorspar cup features vine leaves and grapes carved in low relief (**Fig. 1**, image courtesy of The British Museum).



**Fig. 1** The Barber Cup, located in the British Museum in London

Fluorspar’s beautiful colours and shine granted its permanence in decorative objects from Ancient History throughout the Middle Ages and up until the modern era.

Although fascinating, until the 16<sup>th</sup> century fluorinated minerals were only used for decoration or embellishment, with the exception of Agricola’s use of fluorspar as a flux. History changed for fluorine in the early 1700s, with the efforts of several now well-known scientists, who chased this element despite the practical difficulties of working with it. Glass equipment was excluded and there was danger of severe poisoning.

An interesting use of a fluorine mineral surfaced in 1835, when cryolite glass was used to produce ocular prosthesis.<sup>13</sup>

Dr Lee McDowell reviews these early events in his book ‘Mineral Nutrition History: The Early Years’ (2017).<sup>14</sup> It was the German chemist Andreas Marggraf who took the first step towards fluorine isolation in 1764 by managing to successfully heat fluorite with sulfuric acid to prepare hydrofluoric acid.<sup>15</sup> However, he could not identify the product. Swedish chemist Carl Scheele repeated the procedure in 1771, and he was able to recognise its acidic nature.<sup>16,17</sup>

The dangerous properties of this acid and other fluorinated compounds were later articulated and suffered by Gay-Lussac and Thénard (1809),<sup>18</sup> Rabuteau (1867), Davy and the brothers Knox.<sup>14</sup> Sadly, even more severe consequences befell Paulin Louyet<sup>19</sup> and Jerome Nicklès, who both died from hydrogen fluoride poisoning.<sup>14, 20</sup>

These failed and sometimes dramatic attempts were rewarded in 1813, when Humphry Davy proved that although the substance was analogous to hydrochloric acid, the composing element must be different from chlorine.<sup>21</sup>

Electrolytic processes using hydrogen fluoride were then attempted by Frémy<sup>3</sup> and Gore (1869).<sup>22</sup> Moissan later used those ideas to build his own experiment, for the electrolysis of dry hydrogen fluoride with potassium bifluoride as an electrolyte. This, combined within a Dewar flask, led to the low temperature production of fluorine gas.<sup>1</sup>

The combined efforts of these scientists, sometimes unsuccessful, have significantly changed the world we know today. We now enjoy multiple advantages of adding fluorine to our molecules. The following sections will describe how fluorine has influenced several fields, such as agrochemicals, materials chemistry and the pharmaceutical industry.<sup>23</sup>

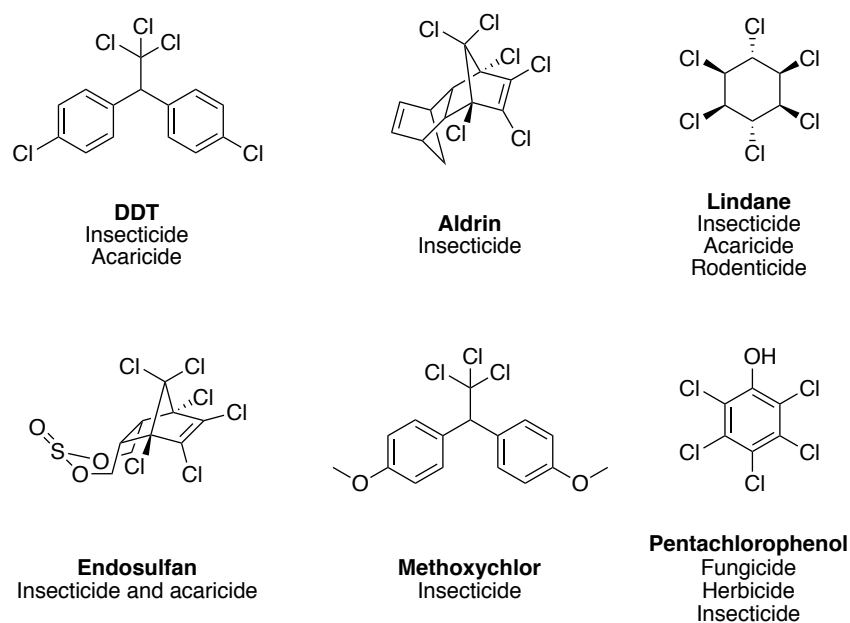
### **1.1.2 Fluorine in agrochemicals**

‘By morning, the wind had brought the locusts; they invaded all Egypt and settled down in every area of the country in great numbers. Never before had there been such a plague of locusts, nor will there ever be again. They covered all the ground until it was black. They devoured all that was left after the hail – everything growing in the fields and the fruit on the trees. Nothing green remained on tree or plant in all the land of Egypt.’ (Exodus 10: 13-15)<sup>24</sup>

Plagues and pests have been a problem for humanity since biblical times, and people had to look for solutions, apart from divine intervention. The first recorded use of insecticides dates back 4500 years, when the Sumerians used sulfur to control the spread of different insects.<sup>25</sup> This and other inorganic elements, like arsenic or lead, have been used continuously throughout history until the 20<sup>th</sup> century. Nowadays, and due to its very high toxicity, sulfur has been generally banned in most countries for this purpose.<sup>25</sup>

With the rapid development of modern science, more complex molecules were synthesised, including those to deter pestilence. Halogenated pesticides were widely used between the 1940s and the 1970s, but they have been banned due to their toxicity and environmental persistence.<sup>26</sup>

Among these halogenated substances, dichlorodiphenyltrichloroethane (DDT) has the highest profile. DDT was used with no official control for decades, but it has extremely damaging effects on wildlife, as well as in human health, due to its capacity to accumulate in adipose tissues.<sup>27</sup> DDT and some other examples of organochlorine pesticides are shown in **Fig. 2**.<sup>28</sup>



**Fig. 2** Organochlorine pesticides from the 20<sup>th</sup> century

The introduction of halogens for pest control continued, even though it still required some effort to reduce their toxicity. The addition of halogens into well-known

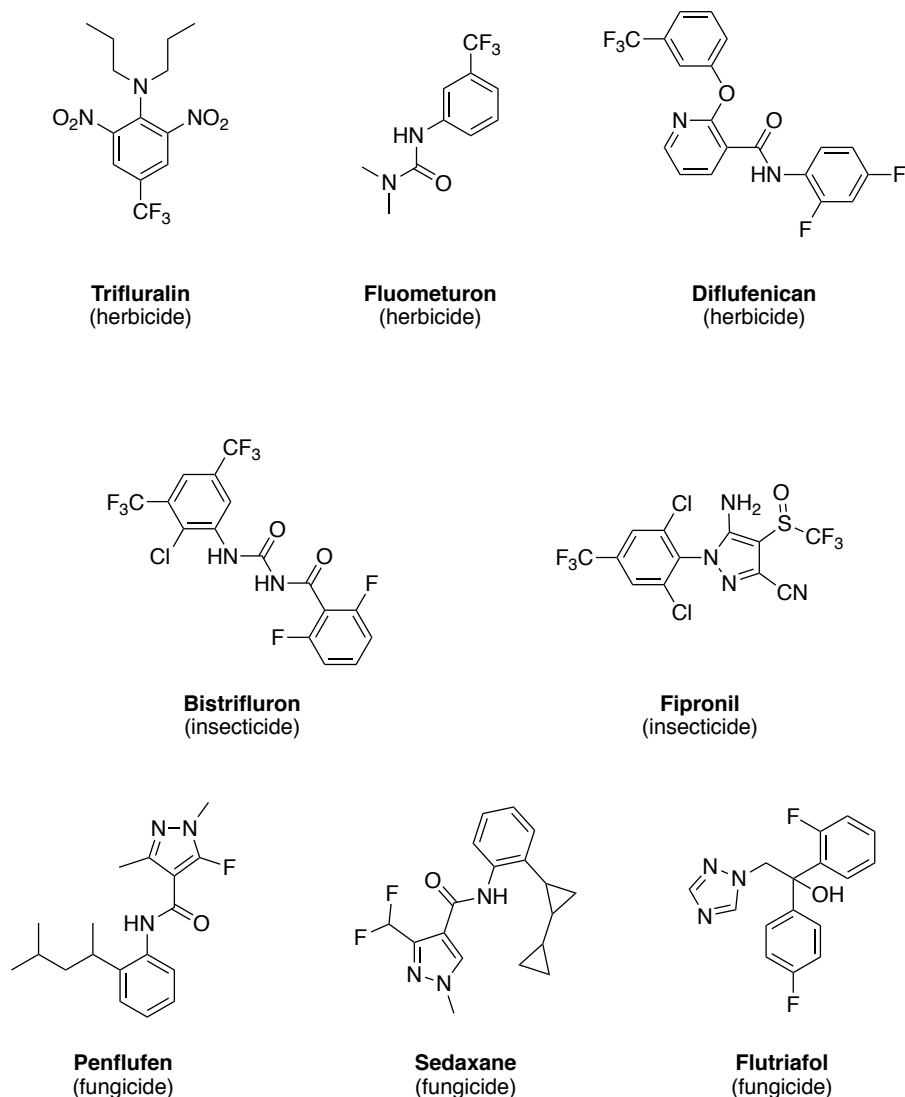
structures by substitution of hydrogen atoms modulates the properties of these products.

Halogen substitution generates a different steric and electronic profile when compared with hydrogen, and such substitutions can change the physicochemical properties of the original product. Thus,  $pK_a$ , lipophilicity, stability and biological activity are all affected when a halogen atom is incorporated into a structure.<sup>29</sup>

Of all the halogens, fluorine has been the most influential in more recent years for agrochemical and medical programmes. Its size – similar to hydrogen's – reduces the steric impact, and being the most electronegative element of the Periodic Table, the effect on the molecule's physicochemical properties can be substantial.

As stated by Theodoridis in 'Fluorine-containing agrochemicals: an overview of recent developments', fluorine can not only dramatically improve the efficacy of the agrochemicals containing it – therefore allowing a lower application dose – but also reduces significantly the environmental impact of these chemicals.<sup>30</sup>

A survey run on modern agrochemicals, provisionally approved by ISO between 1998 and 2008, showed that around 79% contain at least one halogen atom, fluorine being the most common.<sup>29</sup> A pattern has emerged regarding the level of fluorination. Fungicides usually have two fluorine atoms, herbicides at least three, and insecticides often contain a minimum of four fluorine atoms in their structure.<sup>29</sup> A selection of the most successful fluorine-containing pesticides is displayed in **Fig. 3**.



**Fig 3.** Examples of licensed available fluorine-containing pesticides

### 1.1.3 Fluorine in materials chemistry

Like many of the most important discoveries, the introduction of fluorine into materials chemistry was rather serendipitous. In 1938, Plunkett discovered a white powdery residue inside one of his tetrafluoroethylene (TFE) canisters.<sup>31</sup> While he did not know it at the time, he had just found what was to become one of the most enduring non-stick fluorinated materials of all time: Teflon®.

Teflon® (DuPont) is a perfluorinated polymer, prepared from tetrafluoroethylene (PTFE) monomer. It is mainly used in cookery to apply a non-stick cover to pans and pots, and other kitchen utensils. It is extremely resistant to heat variations, and

completely waterproof, properties which make it attractive for a wide range of applications.

PTFE is also used in Goretex® (WL Gore), a protective and waterproof fabric, with properties resembling those of nylon, but with the added advantage of being breathable. Goretex® is formed by a layer of PTFE, surrounded by layers of nylon and polyurethane rendering it gas porous.<sup>32</sup>

Fluorinated polymers have proven to be extremely important in a range of everyday products, but they are also necessary in less mundane fields. An excellent example of this would be Fomblin®, which according to the Solvay website are ‘fluorinated lubricants representing the best choice when aggressive chemical environments, high temperatures or wide-working temperature ranges are involved’.<sup>33</sup> This high performance lubricant is used for machines operating in extreme environments of temperature, such as spacecraft engines or for military purposes.<sup>23</sup>

Fluorine is also a valuable component of liquid crystals, and currently, fluorinated motifs have replaced earlier components such as the cyano group. The advantages they bestow in these materials are many; from improving their reliability to stabilising conformations and rigidify flexible alkyl chains.<sup>34</sup> The location of the fluorine within the liquid crystal enables the tailoring of properties.<sup>35</sup> The field of fluorinated liquid crystals is an expanding one, as personal devices have seen an exponential growth.<sup>36</sup>



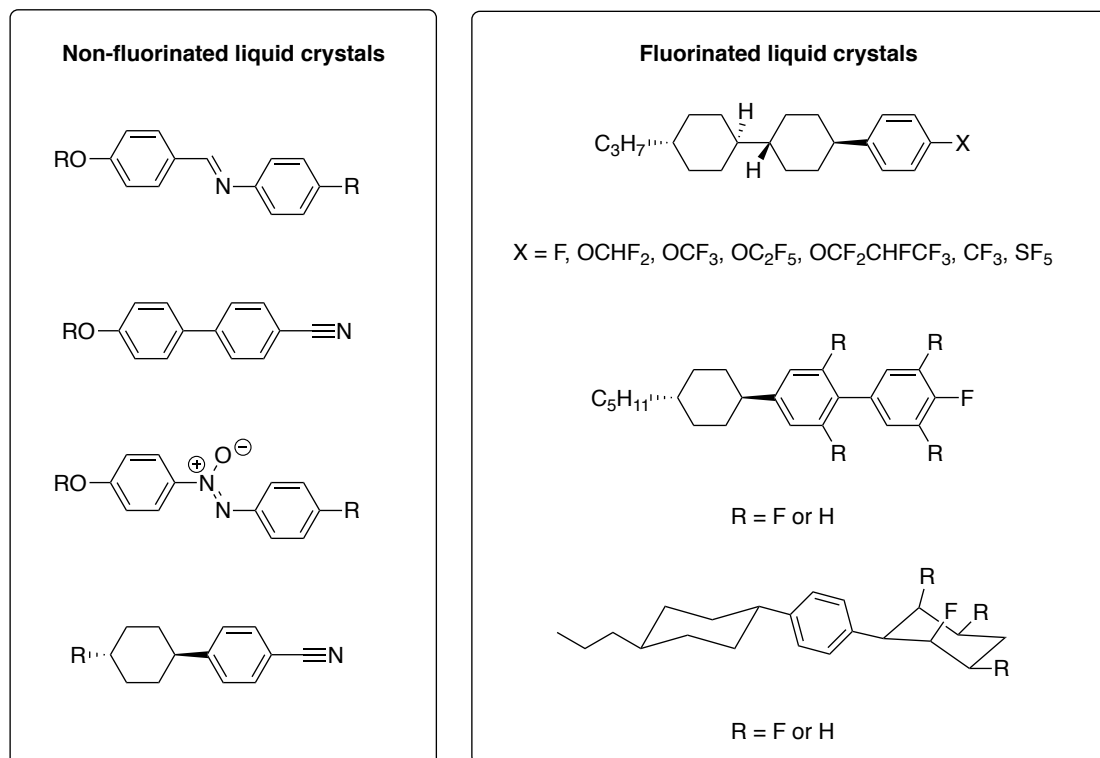


Fig. 4 Successful liquid crystals published in recent years

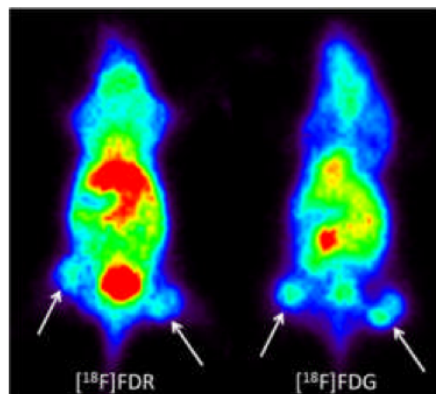
#### 1.1.4 Fluorine in PET

Positron-emission tomography (PET) is a non-invasive technique for clinical imaging of functional metabolic processes in real time.<sup>37</sup> A crucial aspect for a successful PET experiment is the choice of the radioactive isotope, and an important factor is the isotope half-life ( $t_{1/2}$ ).

The most popular positron-emitting isotopes are  $^{15}\text{O}$ ,  $^{13}\text{N}$ ,  $^{11}\text{C}$  and  $^{18}\text{F}$ . Although  $^{15}\text{O}$  is desirable for PET imaging, its half-life of only 2 min is extremely limiting. Slightly longer is the half-life of  $^{13}\text{N}$  (10 min) and  $^{11}\text{C}$  (20 min), but still these can only be used in on-site facilities, since their rapid decay does not allow shipping.<sup>38</sup>

$^{18}\text{F}$  has a much more optimal half-life (109 min). This isotope is most commonly incorporated into fluorodeoxyglucose (FDG), which is taken up by metabolising cells. Furthermore, the sufficiently long half-life of  $^{18}\text{F}$  makes it much more practical for patient preparations, as well as for transport and delivery to a satellite laboratories. [ $^{18}\text{F}$ ]-FDG is used for a wide variety of health areas, including neurology, cardiology and oncology.<sup>39,40</sup>

Other  $^{18}\text{F}$  tracers are being developed as biomarkers. A local example has shown that both enzymatically and chemically prepared  $^{18}\text{F}$ -FDR (fluorodeoxyribose) could be used as an alternative to  $^{18}\text{F}$ -FDG for imaging tumours in mouse models (**Fig. 5**).<sup>40</sup>

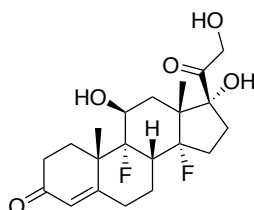


**Fig. 5** [ $^{18}\text{F}$ ]-FDR and [ $^{18}\text{F}$ ]-FDG imaging of tumours in mice models<sup>40</sup>

### 1.1.5 Fluorine in the pharmaceuticals and medicinal chemistry

The substitution of hydrogen by fluorine has proven popular for tuning the properties of bioactive compounds, and in this context, fluorine applications continue to increase in importance in the pharmaceutical industry.<sup>41</sup>

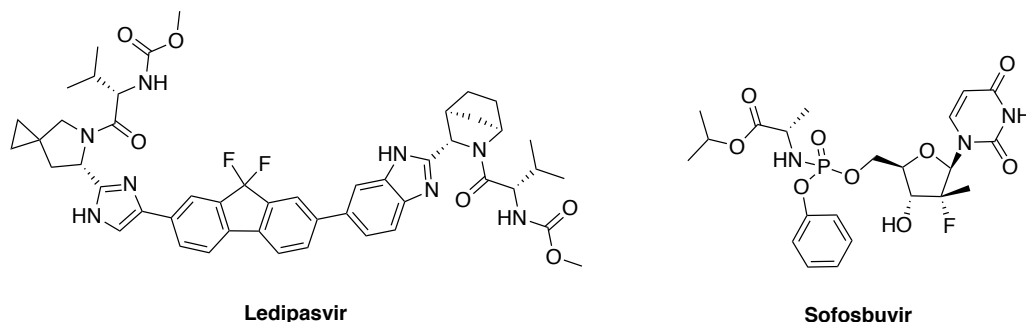
In 1955, the U.S. Food and Drug Administration (FDA) approved the first ever fluorine-containing drug.<sup>42</sup> This was the steroid fludrocortisone; which is used to treat low glucocorticoid levels caused by diseases of the adrenal gland, such as Addison's disease. Fludrocortisone (**Fig. 6**) was developed from a glucocorticoid, a natural steroid generated in the human body.<sup>43,44</sup>



**Fig. 6** Structure of first fluorinated drug approved, fludrocortisone

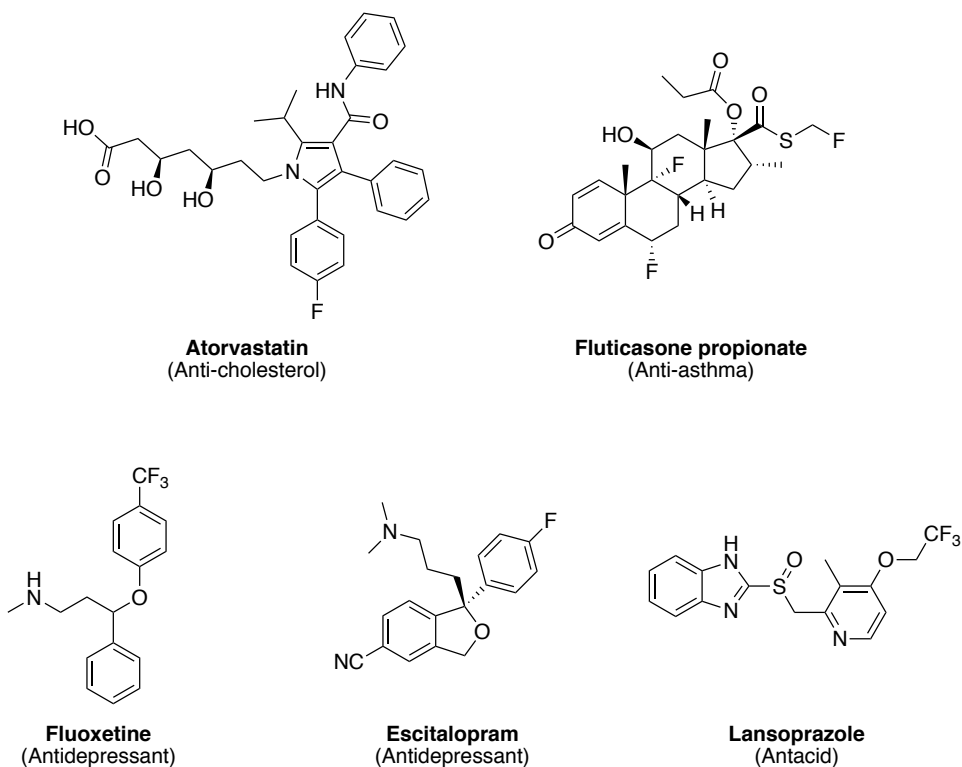
Between 2008 and 2010, it was suggested that fluorine was present in at least 20% of commercial drugs.<sup>45</sup> However by 2016 this had grown to about a 30%.<sup>42</sup>

The current top-selling drug is a combination of ledipasvir-sofosbuvir (trade name Harvoni), used for the treatment of hepatitis C. In 2015, the sales from this drug totalled \$14.3 billion!<sup>46</sup>



**Fig. 7** Structure of Ledipasvir and Sofosbuvir, for treatment of hepatitis C

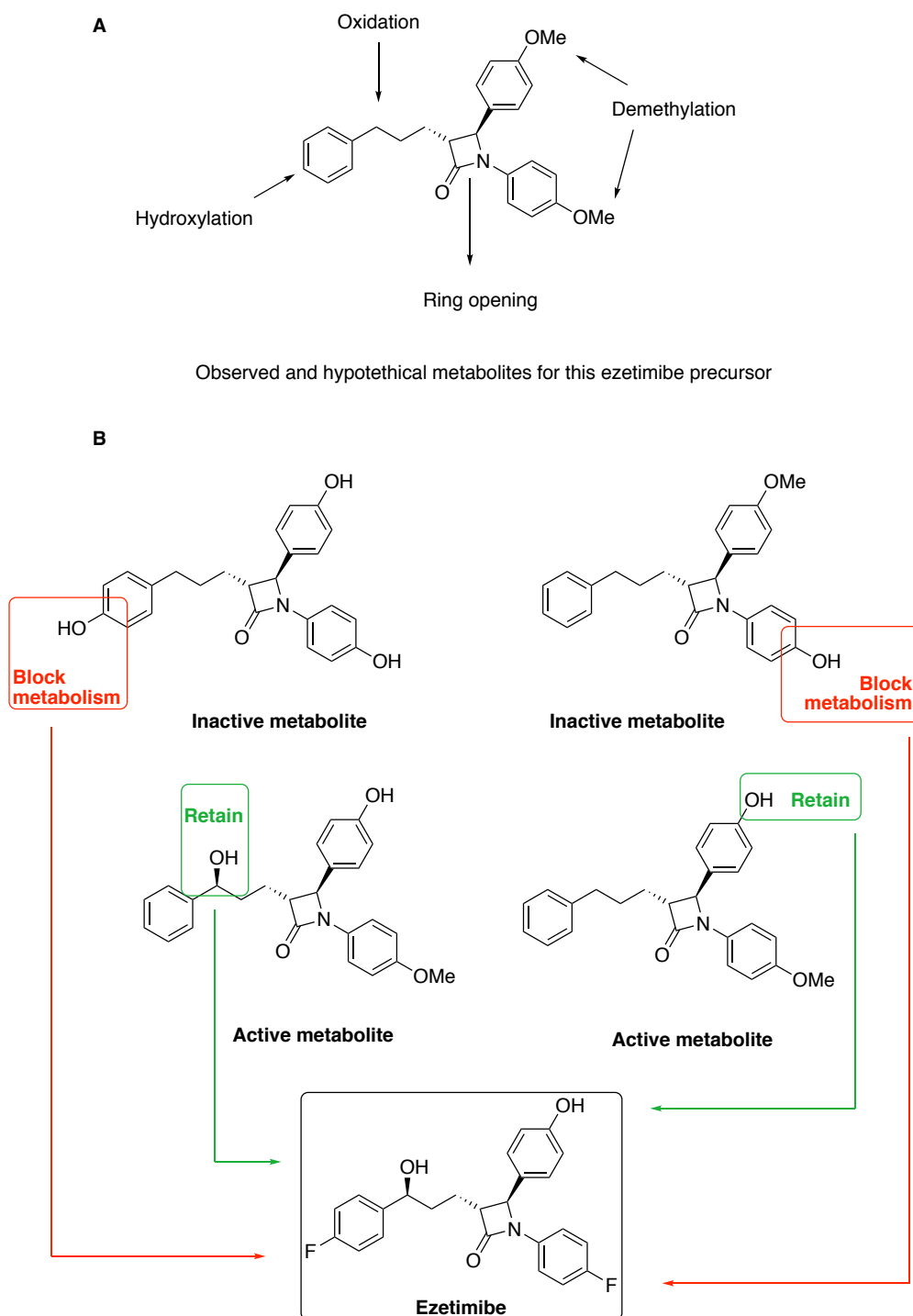
Sofosbuvir took over from atorvastatin (trade name Lipitor) – a fluorine-containing cholesterol biosynthesis inhibitor – as the top selling drug. This was ranked the number one best-selling drug for the previous decade. It was closely followed by the fluorine-containing drug fluticasone propionate (Advair Diskus, an anti-asthma drug) at number four in 2010, and lansoprazole (or Prevacid, an antacid remedy) at number five. Escitalopram (Lexapro), located at number eleven has effectively replaced the CF<sub>3</sub>-containing fluoxetine (Prozac) as the leading antidepressant. **Fig. 8** displays these structures.<sup>47</sup>



**Fig. 8** Fluorinated blockbusters in medicinal chemistry

It is evident that fluorine provides some refined properties to these molecules. Among those improvements are metabolic stability and bioavailability or enhanced protein-ligand interactions. Other properties affected include lipophilicity and  $pK_a$ .<sup>48</sup>

Metabolic stability constitutes a crucial factor when designing a drug, since it determines the drug's efficacy. If a molecule is too labile and prone to be readily decomposed, metabolic oxidation processes will be fast and abundant, facilitating the drug's excretion prior to its effect. This fast elimination of such a xenobiotic from the body will reduce its efficacy. Strategic fluorination has become one of the most effective solutions for this problem, leading to a generation of drugs who are up to 400 times more potent at lower doses due to their increased *in vivo* metabolic stability.<sup>49</sup> For example, ezetimibe (a cholesterol reducing drug) has been optimised to its current structure through several improvements, as shown in **Scheme 1**.<sup>50, 51, 52, 53</sup>



**Scheme 1.** Optimisation process of ezetimibe<sup>53</sup>

In ezetimibe's case study, **Scheme 1a** displays the major metabolic issues with the drug's precursor, while **Scheme 1b** shows how they were solved, leading to the active entity that is sold today.<sup>53</sup>

A second factor that affects efficacy is lipophilicity. Lipophilicity determines the permeability of the drug through the cell membrane and its interactions with proteins such as albumin, and it is a property that requires careful tuning in medicinal chemistry development. A lipophilicity that is too high will result in the molecule being trapped within the lipid bilayer or bound to protein; if it is too low the product will remain outside of the cell and be excreted by the urinary system.

It has often been assumed that fluorine incorporation increases lipophilicity, however, this statement needs to be refined, according to the degree or type of fluorination. Mono- to trifluorination of saturated alkyl groups tend to decrease lipophilicity, whereas fluorination of aromatic or heterocycle rings tends to increase lipophilicity.<sup>54,55</sup>

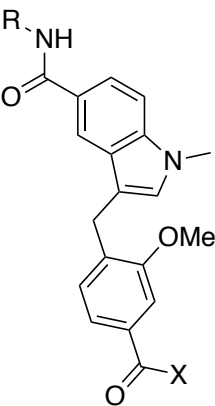

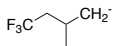

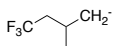

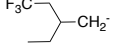

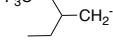
The influence of monofluorination on Log *P* values is reflected in **Table 1**, an example presented by Smart in 2001.<sup>56</sup> In this case, the effect of fluorine substitution is described for small aliphatic molecules, where the Log *P* values immediately show a substantial decrease.

**Table 1.** Effect of monofluorination on Log *P* in small aliphatic molecules

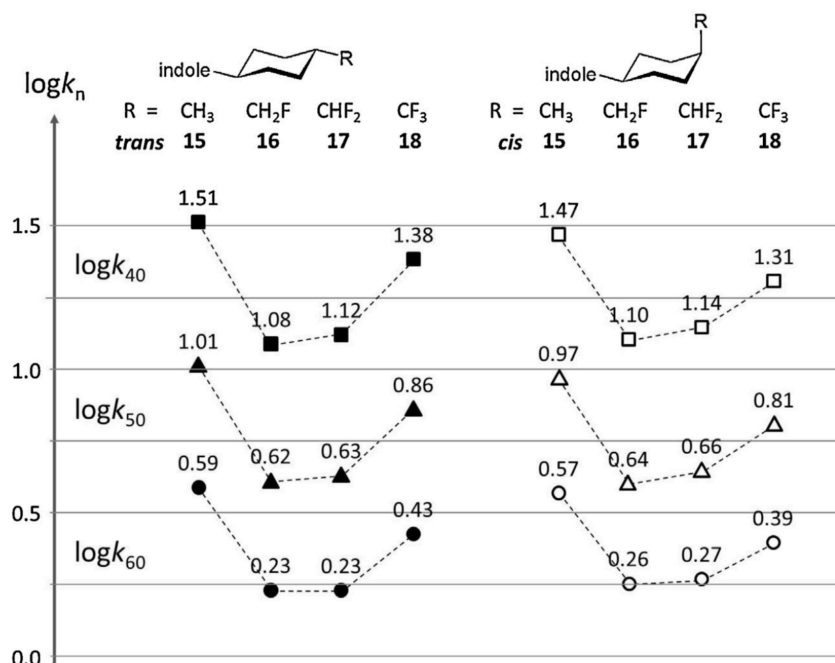
Molecule	Log <i>P</i>
CH <sub>3</sub> CH <sub>3</sub>	1.81
CH <sub>3</sub> CH <sub>2</sub> F	0.75
CH <sub>3</sub> (CH <sub>2</sub> ) <sub>3</sub> CH <sub>3</sub>	3.11
CH <sub>3</sub> (CH <sub>2</sub> ) <sub>3</sub> CH <sub>2</sub> F	2.33

**Table 2** shows the influence of trifluoromethylation in different substituents for leukotriene antagonists, as reported by Jacobs *et al.* in 1994.<sup>57, 58</sup> For this more complex molecule, trifluoromethylation causes the lipophilicity values to increase, in comparison with the non-fluorinated analogues.

**Table 2.** Effect of trifluoromethylation on leukotriene antagonists<sup>57,58</sup>

Structure	R	X	Measured Log <i>P</i>
		<i>o</i> -SO <sub>2</sub> Tol	5.85
		<i>o</i> -SO <sub>2</sub> Tol	6.18
		OH	4.86
		OH	5.12
		<i>o</i> -SO <sub>2</sub> Tol	6.29
		<i>o</i> -SO <sub>2</sub> Tol	6.45
		OH	5.42
		OH	5.53

Müller and co-workers have carried out multiple studies on how fluorine modifies the partition coefficient. **Fig. 9** shows how the logarithm of the capacity factor (Log *k*), directly related to the Log *P*, behaves under the different degrees of fluorination for two series of diastereomeric indoles, with varying substituents. The image also depicts the changes in three sets of solvent conditions: 40% acetonitrile in water (log *k*<sub>40</sub>), 50% of acetonitrile (log *k*<sub>50</sub>) in water and 60% of acetonitrile in water (log *k*<sub>60</sub>).<sup>59</sup>



**Fig. 9** Fluorination effect on the log of the capacity factor (*k*) (lipophilicity) for two series of indoles

The discussion on lipophilicity is a continuous one in medicinal chemistry. Historically, Lipinski defined the desired properties for a drug candidate in the commonly known ‘Rule of Five’. This popular rule explained that the perfect candidate would possess less than five hydrogen bond donors and ten hydrogen bond acceptors, a molecular mass of less than 500 Da, and finally, a partition coefficient factor (Log *P*) lower than five.<sup>60</sup> Generally, pharmaceutical companies look for Log *P*s lower than three.<sup>61</sup>

Drugs targeting the Central Nervous System should have lipophilicities not higher than 2.5 (but ideally on the 1.5 – 2.5 range);<sup>62,63,64</sup> while for drugs orally or intestinally absorbed, ideal values are between 1.35 – 1.8.<sup>65</sup> Sublingual drugs can be used with a Log *P* closer to 5.<sup>66,67</sup>

Acidity and basicity can also be modulated by the introduction of fluorine atoms. This is usually measured by p*K*<sub>a</sub>.<sup>68</sup> The addition of fluorine into non-fluorinated structures usually decreases the p*K*<sub>a</sub> values, thus making the structures more acidic. **Table 3** displays p*K*<sub>a</sub> values for a series of carboxylic acids, alcohols and amines.<sup>68</sup>



**Table 3.** Fluorination effect on pK<sub>a</sub>

Carboxylic acid	pK <sub>a</sub>	Alcohol	pK <sub>a</sub>	Amine	pK <sub>a</sub>
CH <sub>3</sub> CO <sub>2</sub> H	4.8	CH <sub>3</sub> CH <sub>2</sub> OH	15.9	CH <sub>3</sub> CH <sub>2</sub> NH <sub>2</sub>	10.7
CH <sub>2</sub> FCO <sub>2</sub> H	2.6	CF <sub>3</sub> CH <sub>2</sub> OH	12.4	CH <sub>2</sub> FCH <sub>2</sub> NH <sub>2</sub>	9.0
CHF <sub>2</sub> CO <sub>2</sub> H	1.3	(CH <sub>3</sub> ) <sub>2</sub> CHOH	17.1	CHF <sub>2</sub> CH <sub>2</sub> NH <sub>2</sub>	7.3
CF <sub>3</sub> CO <sub>2</sub> H	0.5	(CF <sub>3</sub> ) <sub>2</sub> CHOH	9.3	CF <sub>3</sub> CH <sub>2</sub> NH <sub>2</sub>	5.7
CH <sub>3</sub> CH <sub>2</sub> CO <sub>2</sub> H	4.9	(CH <sub>3</sub> ) <sub>3</sub> COH	19.0		
CF <sub>3</sub> CH <sub>2</sub> CO <sub>2</sub> H	3.1	(CF <sub>3</sub> ) <sub>3</sub> COH	5.4		
C <sub>6</sub> H <sub>5</sub> CO <sub>2</sub> H	4.2	C <sub>6</sub> H <sub>5</sub> OH	10.0		
C <sub>6</sub> F <sub>5</sub> CO <sub>2</sub> H	1.7	C <sub>6</sub> F <sub>5</sub> OH	5.5		

## 1.2 Metabolism of Drugs

Drug metabolism is closely associated with the pharmacokinetic properties displayed by a molecule. Absorption, distribution, metabolism and excretion (ADME) processes are extremely important aspects in drug discovery, since they determine the behaviour of a certain candidate within the human body. The development of a new drug and the optimisation of the ADME properties takes, on average, over \$2.6 billion and more than a decade to be completed.<sup>69,70</sup>

The cost and time spent usually include multiple failed attempts, detours, and fresh starts. In the whole, the design of new commercial drugs is a complex procedure, and special care must be taken when the outcome affects living organisms.

In humans, most of the enzymatic drug metabolism is carried out by the liver, which is the most active tissue for processing xenobiotics. However, some drugs that are not metabolised by the liver can be degraded in the kidneys, intestines – including gut

flora – or lungs. In rare occasions, drugs can be metabolised in other tissues like the brain, heart, muscle, spleen or even the skin.<sup>71</sup>

The metabolism of a drug can be viewed to progress through three stages: Phase I, Phase II and Phase III. Drugs are xenobiotics so an organism will operate to eliminate any traces of them from its system.

Phase I involves modification of the molecule through oxidation, reduction or hydrolysis. Groups such as alcohols, carboxylic acids, amines or thiols are added to the original molecule.<sup>71</sup> In particular, oxidation is usually carried out by the cytochrome P<sub>450</sub> (CYP) enzymes, which hydroxylate and make the xenobiotic more soluble, a process that will be described in more detail in **Section 1.3.2**.

Phase II or the conjugation phase – requires further modification, usually by the attachment (conjugation) of an ionised (eg. sulfate) or polar (eg. glucose) group that will enhance the water solubility of the metabolite for its excretion by the kidneys.

The aim of the reactions occurring during Phase II is to inactivate the remaining drugs, or any of their active phase I metabolites.<sup>71</sup> The transformations are also meant to produce a much more water-soluble derivative of the xenobiotic, and include the generation of glucuronides, sulfates, or the introduction of glycine, glutamine or cysteine.<sup>71</sup> All of these changes then lead to Phase III, which involves excretion in the urine, bile or sweat.

Pharmaceutical companies and research groups generally avoid experimentation with higher organisms until necessary. Cost increases considerably, and there are ethical and moral challenges with animal experimentation. Therefore, the use of more simple and readily available microbial models is attractive, certainly at early stage investigations. One of these microbial models is the fungus *Cunninghamella elegans*.

### **1.3 *Cunninghamella elegans* as a model for mammalian drug metabolism**

#### **1.3.1 The fungus *Cunninghamella elegans***

The genus *Cunninghamella* was described by Matruchot in 1903.<sup>72</sup> *Cunninghamella* is a filamentous fungus that is found mainly in soil and plants, in Mediterranean and subtropical zones.<sup>73</sup> This genus is currently formed by 14 species, some of which have been reported as human mucormycosis.<sup>72</sup> Particularly, *C. bertholletiae* has been classified as a human and animal pathogen.<sup>73,74</sup>

Initially, up to three different studies (Samson in 1969,<sup>75</sup> Baijal and Mehrotra in 1980<sup>76</sup> and Lunn and Shipton in 1983<sup>77</sup>) described *C. bertholletiae* and *C. elegans* as the same species, given that there were no significant differences in their morphological characteristics.<sup>72</sup> However, Weitzman and Crist (1979)<sup>78</sup> managed to make a distinction between both species according to their growth capacity under different temperatures. The maximum temperature of growth for *C. elegans* occurs at 37 °C, while *C. bertholletiae* can grow at temperatures between 40 – 45 °C.<sup>72</sup> In addition to that debate, Weitzman made a case in 1984 as to why *C. elegans*, *C. bertholletiae* and *C. echinulata* should be maintained as three different species. The argument is based on the fact that those three fungi display different mating behaviour, morphology, and temperature tolerance, amongst other factors.<sup>79</sup>

*C. elegans* has very mild culture conditions and fast growth; it matures after a 72 h incubation at 28 °C. The appearance of the mature fungus in agar plates is cotton-like, of a very pale grey or yellowy tint, where the filaments are easily observed by the naked eye. When shaken in the liquid medium, the filamentous *C. elegans* tends to form spherical pellets. However, whenever bacterial contamination is present, the culture fluid quickly becomes turbid, and the hyphal pellets are not formed as usual.<sup>80</sup>



**Fig. 10** Agar plates of *Cunninghamella elegans*

*C. elegans*, like most of the *Cunninghamella* sp., possesses cytochrome P<sub>450</sub> enzymes, which are analogous to those in mammals. This fungus is able to metabolise a wide range of xenobiotics using phase I oxidative and phase II conjugative biotransformations,<sup>73</sup> and therefore, *C. elegans* has found use as an animal model.

### 1.3.2 Cytochrome P<sub>450</sub> and metabolism

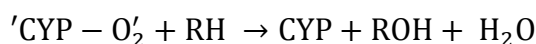
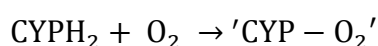
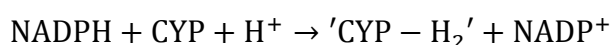
Drugs tend to be metabolised in such a way that excretion from the organism becomes possible. Cytochrome P<sub>450</sub> enzymes play a role in phase I metabolism.

In humans, P<sub>450</sub> enzymes reside principally in the liver, but can also be found in the lungs and intestines, amongst other organs.<sup>71</sup> In particular, they predominantly exist in the endoplasmic reticulum of the hepatocytes.<sup>73</sup> Cytochrome P<sub>450</sub> (CYP) is a carbon monoxide binding haemoprotein<sup>71</sup> mainly involved in oxidative reactions, including carbon hydroxylation, heteroatom and/or double bond oxidation. However, in certain cases, reports of reductive biotransformations by these enzymes have also been described.<sup>73</sup>

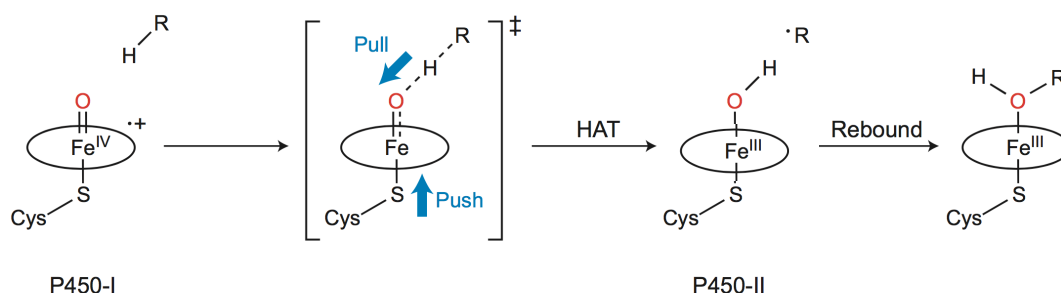
The way cytochrome P<sub>450</sub> enzymes work usually requires an initial binding of the drug to the active site of the enzyme, before it can take part in the 'P<sub>450</sub> catalytic cycle'.<sup>73</sup> It is largely accepted that they possess an open active site, and usually prefer lipophilic substrates.<sup>73</sup> The quality of binding, as well as the specific orientation of a molecule binding CYP, largely depends on hydrophobic and steric interactions between the substrate and the enzyme; and these interactions are highly dependant on lipophilicity.<sup>73</sup>

The enzymes from this complex do not work independently, and they require the electron-donating activity of a second enzymatic system, in order to remain catalytically active. In mitochondria or prokaryotes this is a system of two proteins, while for the endoplasmic reticulum enzymes, the electrons are donated by just one other enzyme: cytochrome P<sub>450</sub> reductase (or CPR).<sup>73</sup> Additionally, CYP enzymes need NADPH as a coenzyme, and oxygen as a substrate, in order to function.<sup>73</sup>

Oxidation by cytochrome P<sub>450</sub> starts with hydride reduction (2 electron transfer) by NADPH and protons from the medium. The reduced and protonated CYP ('CYP-H<sub>2</sub>') then binds molecular oxygen, and the active complex formed ('CYP-O<sub>2</sub>') is now able to oxidise the drug (RH), releasing water and regenerating the original CYP.<sup>71</sup>



However, a recent study has shown that the 'CYP-O<sub>2</sub>' complex, in which the iron (IV) is linked to the molecular oxygen, had a basic pH.<sup>81</sup> This result indicated that the 'CYP-O<sub>2</sub>' species was not the previously presumed iron-oxo Cys-S-Fe(IV)=O, but it was the hydroxy form Cys-S-Fe(IV)-OH instead (**Scheme 2**).<sup>81,82</sup>

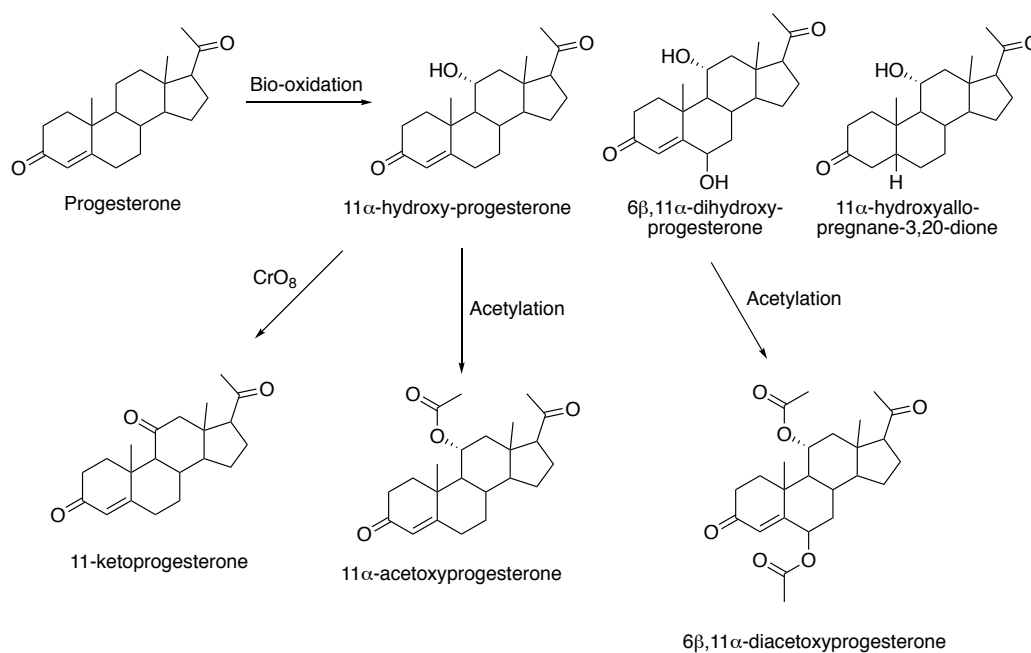


**Scheme 2.** Pull-push mechanism in cytochrome P<sub>450</sub>.<sup>82</sup>

The basicity of that oxygen has been explained by the electron 'push' generated by the axial cysteine-thiolate ligand.<sup>81,82</sup> This 'electron push' from the sulphurated ligand creates a 'pull' from the ferryl oxygen that allows the C-H bond rupture (**Scheme 2**).<sup>82</sup> This balanced force from the ligands' bonds to iron recreates a pull-push scenario that has been known as the 'pull-push' mechanism in cytochrome P<sub>450</sub>.

### 1.3.3 Biotransformation of xenobiotics

In 1952, Peterson and Murray pioneered in the use of microbial biotechnology for steroid metabolism with the biotransformations of progesterone by a *Rhizopus* organism (**Scheme 3**).<sup>83</sup> They demonstrated the use of microbiological tools for the synthesis – and scale-up – of otherwise chemically challenging transformations, by exploiting the natural metabolic pathways of different microorganisms.



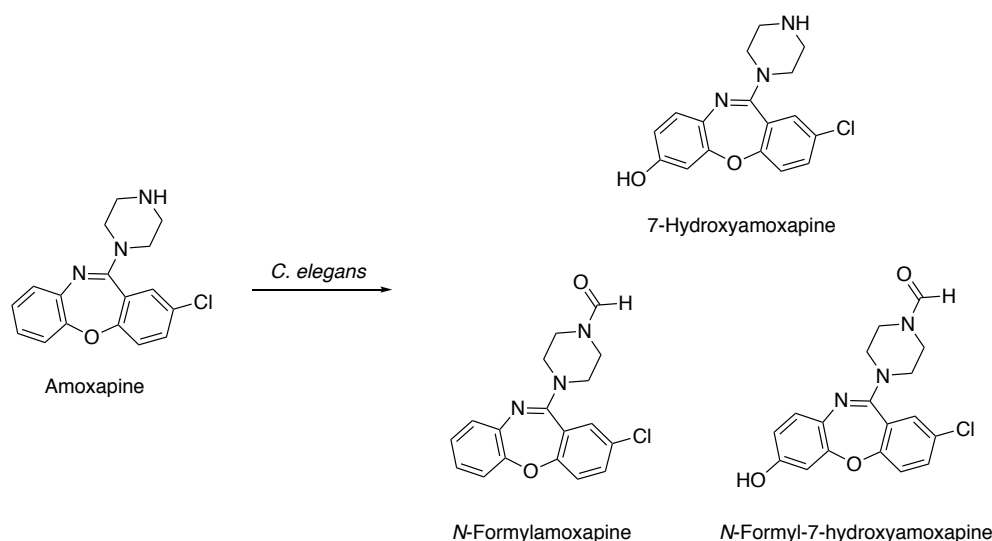
**Scheme 3.** Experiments by Peterson and Murray, using biotransformations to facilitate further reactions<sup>83</sup>

Nevertheless, the use of microbes as simplified models for mammalian drug metabolism did not appear for almost two decades after these seminal studies. Beukers *et al.* published a book in 1972 entitled ‘Microbial conversion as a tool in preparation of drugs’ where they studied the application of microorganisms to assist in the discovery of new drugs.<sup>84</sup>

*Cunninghamella* began to be used as a metabolic model in 1974.<sup>85</sup> One of the early examples include the stereoselective metabolism of warfarin by *C. elegans*.<sup>86</sup>

Cerniglia in particular has made a significant contribution to metabolism studies with *Cunninghamella elegans*, over 40 years. Some of his latter contributions with this organism include the biotransformation of the antidepressant amoxapine (2000)<sup>87</sup> and

malachite green (2001)<sup>88</sup>. The hydroxylation and formylation of amoxapine by *C. elegans* is illustrated in **Scheme 4**.<sup>87</sup>

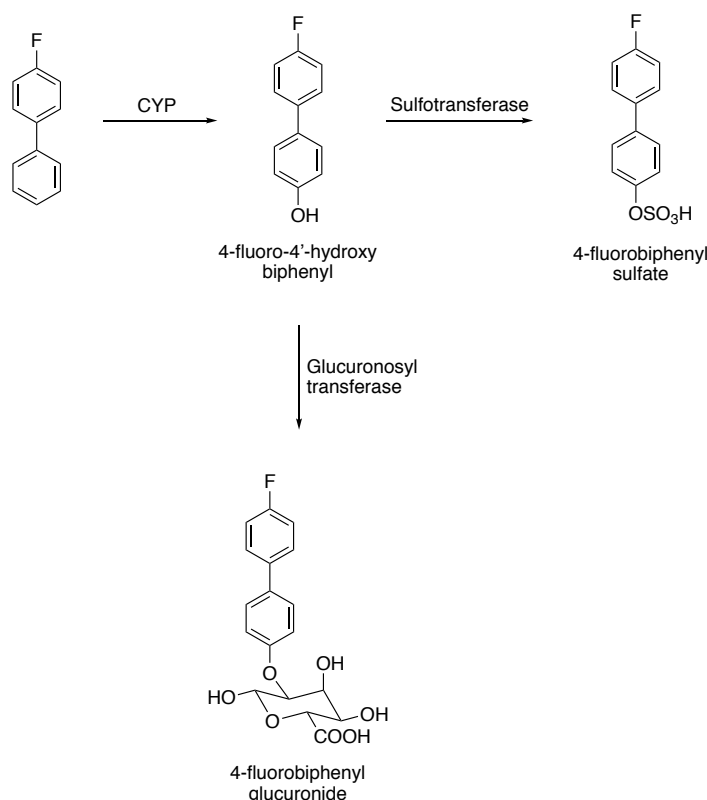


**Scheme 4.** Biotransformation of amoxapine in *C. elegans*<sup>87</sup>

*C. elegans* has also been used to explore the metabolism of fluorinated drugs. Within this topic, Murphy compared different strains of *Cunninghamella* and bacterial strains. For example, the biotransformation of the anti-cancer drug flutamide in *C. elegans*, *C. echinulata* and *C. blakesleeana* has been evaluated. In that case, *C. elegans* DSM 1908 produced the best profile, affording phase I metabolites, and was chosen for more detailed analysis.<sup>89</sup>

During the biodegradation of the anti-inflammatory flurbiprofen with *C. elegans* and *C. blakesleeana*, some phase II conjugated metabolites were observed, showing that the fungus can progress beyond phase I metabolites in some cases.<sup>90</sup>

Another example focused on the biotransformation of 4-fluorobiphenyl in *C. elegans* which demonstrated phase I mono- and di-hydroxylation products, the major being 4-fluoro-4'-hydroxybiphenyl (**Scheme 5**). The fact that there are multiple hydroxylations, suggests that *C. elegans* is a useful model for mammalian drug metabolism and as a means to prepare candidate metabolites.<sup>91</sup>



**Scheme 5.** Biotransformation of 4-fluorobiphenyl by *Cunninghamella elegans*<sup>91</sup>

In 2013, Murphy and co-workers explored biotransformations in *C. elegans* biofilms, in place of the traditional suspension cultures. Fungal biofilms present the advantage of being easily recyclable, and have an increased facility of handling.<sup>92</sup> Several studies have been published since then with biofilms as an improved approach for metabolism investigations. Both synthetic pharmaceuticals<sup>93</sup> and relatively new motifs – such as sulfanylpentafluorides<sup>94</sup> – were explored.

These fungal biofilms studies allowed screening, as well as an ease in the scale-up of fluorodrug metabolites for characterisation.<sup>95</sup> Many compounds were not easily accessible by chemical methods. The mixture of both a chemical and a biological approach is extremely fruitful, and is being used increasingly to access fluorinated intermediates or for the easier production of fluorometabolites.<sup>96, 97</sup>

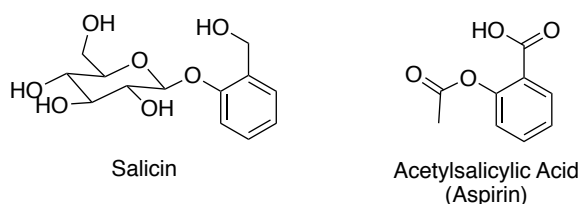


## 1.4 Natural Products. Relevance and biosynthesis

### 1.4.1 Relevant natural products. Discovery, roles and applications

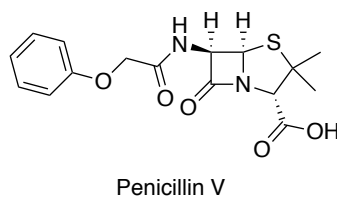
Nature has played a crucial role in providing structurally complex chemical compounds. And many natural products are found to be bioactive. However, the isolation of these products can be difficult and chemists have placed a great effort into their total synthesis.

In 1763, Stone described how the bark of willow trees had powerful astringent properties, and related them to the presence of salicylates (a general term used for the derivatives of salicylic acid).<sup>98,99</sup> Although Stone was the first person to relate the medicinal properties of the willow bark to the salicylates, he was not the first to report its properties. Mention to the curative and pain-relieving effects of willow tree's bark were already catalogued 2500 years previous, by Greek physician Hippocrates, and continued throughout the years of the Roman Empire.<sup>99</sup> Acetylsalicylic acid is the active principle of aspirin, as we know it today.



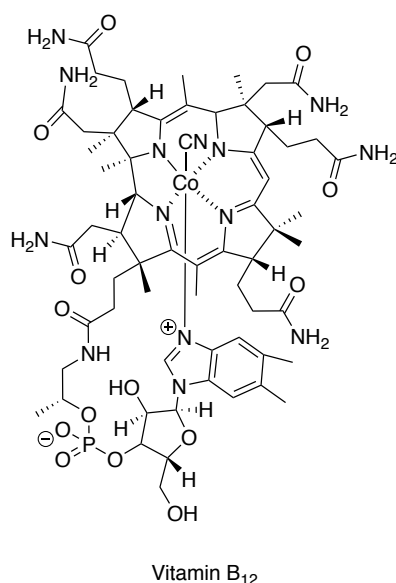
**Fig. 11** Structures for salicin (a salicylate derivative) and acetylsalicylic acid

In 1928, Alexander Fleming noticed that one of his agar plates – initially prepared to grow staphylococcus – had grown a mould impurity. After realising that the mould had effectively killed the surrounding bacteria, he decided to cultivate the new microorganism separately.<sup>99,100</sup> This new organism was later found to be *Penicillium notatum*, and Fleming named the powerful antibacterial agent isolated from it as ‘penicillin’ (**Fig. 12**).<sup>99,100</sup> The isolation and medicinal properties of penicillin were subsequently demonstrated by a team led by Florey and Chain and in 1945, Fleming, Florey and Chain were awarded the Nobel Prize in Physiology or Medicine for ‘the discovery of penicillin and its curative effect in various infectious diseases’.<sup>101</sup>



**Fig. 12** Structure of penicillin, as isolated by Fleming from *Penicillium notatum*

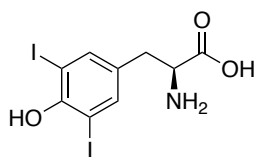
Cobalamin, also known as vitamin B<sub>12</sub>, is a water-soluble vitamin – and a source of the primary function of cobalt in humans. This vitamin acts as a cofactor of two crucial enzymes for human's health (methionine synthase and methylmalonyl-CoA mutase), and it also has a role as a transporting system.<sup>102</sup> The structure of vitamin B<sub>12</sub> (**Fig. 13**) was solved by Dorothy Crowfoot-Hodgkin,<sup>103</sup> in 1956 by X-ray crystallography, in what has been described as 'one of the finest contributions of British science'.<sup>99,104</sup> Crowfoot-Hodgkin was awarded the Nobel Prize in Chemistry in 1964 for 'her determinations by X-ray techniques of the structures of important biochemical substances', including penicillin.<sup>105</sup>



**Fig. 13** Structure of Cobalamin, by Dorothy Crowfoot-Hodgkin

#### 1.4.2 Halogenated natural products and enzymatic halogenation

The first halogenated natural product was isolated and identified at the end of the 19<sup>th</sup> century by Drechsel. Investigations with the coral *Gorgonia cavolii* led to the discovery of diiodotyrosine (**Fig. 14**).<sup>106</sup>

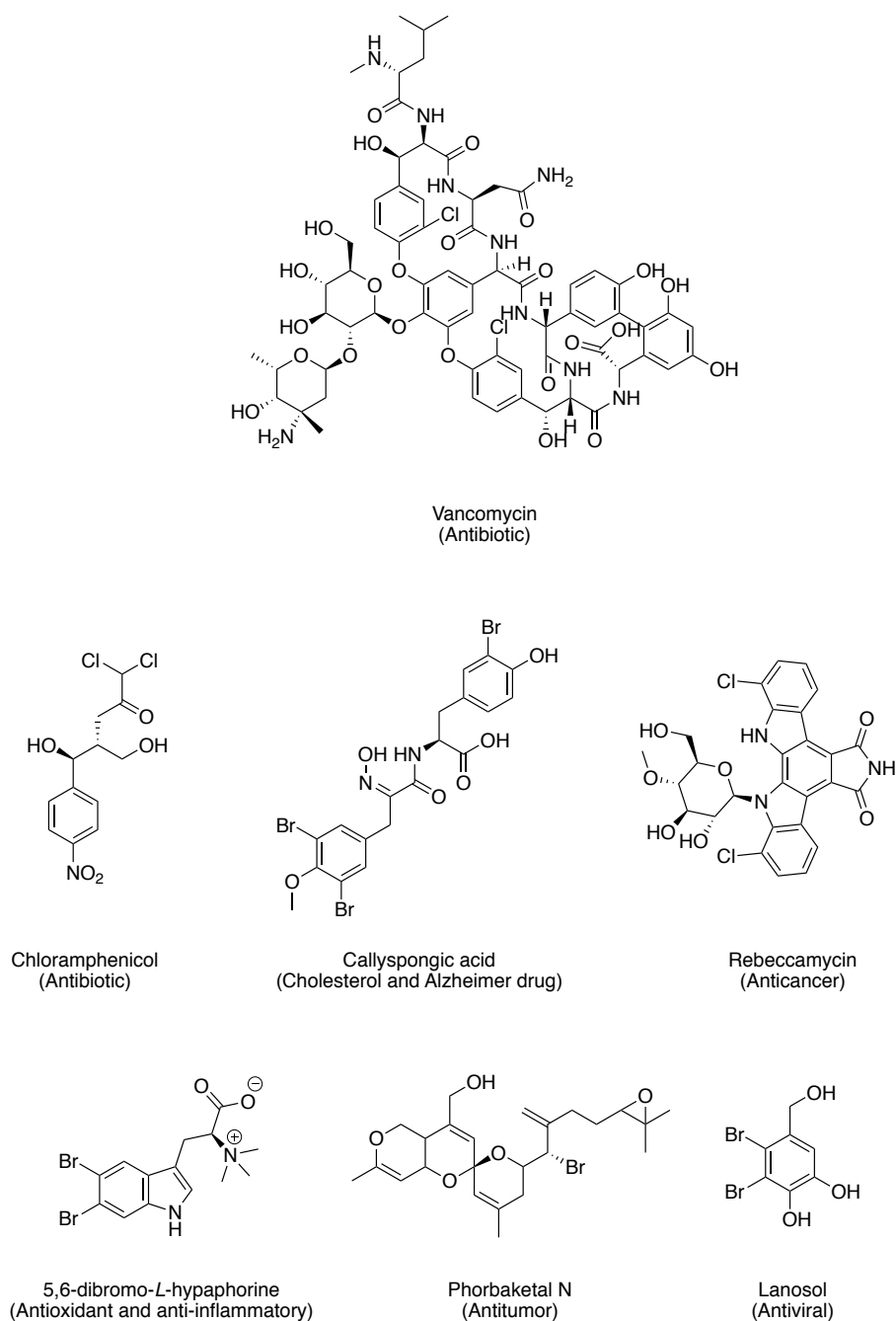


Diiodotyrosine

**Fig. 14** Diiodotyrosine, the first halogenated natural product to be discovered

Since then, the number of halogenated natural products has grown considerably. In 1973, a review stated that more than 200 halogenated products had been isolated from natural sources; while in 1991 that number had grown to 1500.<sup>107</sup> By the end of 2010, more than 4700 halogen-containing structures from nature had been reported.<sup>108</sup> More than 5000, were identified by 2015, the majority occurring in marine sources.<sup>109</sup> Chlorinated products are the most abundant, followed by their brominated analogues. However, only around a hundred feature iodine, and fluorinated natural structures are extremely rare.<sup>108</sup>

The presence of the halogen atoms in the structure of a natural product has a marked effect on its properties, and actively influences biological activity.<sup>110</sup> The bioactivities displayed are varied. Some halogenated antibacterial, antifungal and antiviral drugs include the well-known vancomycin, chloramphenicol or lanosol. One of the most sought-after drugs are antitumor and anticancer agents, such as rebeccamycin or phorbaketol N. 5,6-Dibromo-*L*-hipaphorine is an antioxidant and anti-inflammatory product isolated from the sponge *Hyrtios* sp., indigenous to Fiji. Compounds displaying enzymatic and molecular activity are also on the halogenated spectrum; the brominated callispongic acid is a bromotyrosine that mediates cholesterol metabolism and is of use in the treatment of Alzheimer's disease.<sup>110,108</sup> These structures are shown in **Fig. 15**.

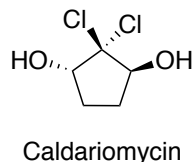


**Fig. 15** Halogenated natural products and their biological activity

The biosynthetic pathways to some of these metabolites has been studied as they hold the key to facilitate the discovery of new halogenated natural products.<sup>109</sup>

The first halogenating enzyme (halogenase) was  $\beta$ -ketoadipate chlorinase, reported by Hager and co-workers in 1959.<sup>111</sup> This enzyme was found while investigating the biosynthesis of caldariomycin (**Fig. 16**) – a di-iodinated cyclopentane isolated from *Caldariomyces fumago*. This chlorinase is a type of haloperoxidase, able to catalyse

the C-Cl bond formation of the metabolites from the caldariomycin metabolic pathway, whenever hydrogen peroxide and chloride ions are present in the biosynthetic medium.<sup>111</sup>

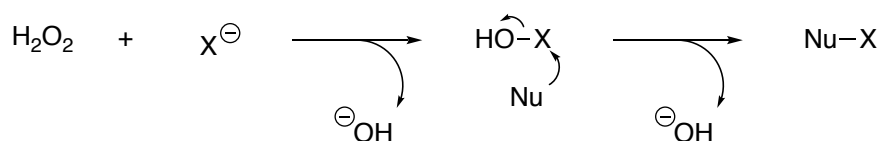


**Fig. 16** Structure of caldariomycin

Halogenases are classified according to several parameters – type of halogen, origin or functionality, among others or according to their mechanism, which span nucleophilic, electrophilic or radical.<sup>112</sup>

$\beta$ -Ketoadipate chlorinase is a haloperoxidase.<sup>111</sup> Haloperoxidases are the first of the two major classes of enzymes involved in electrophilic halogenation, the other being flavin-dependent halogenases.<sup>112</sup>

The exact mechanistic pathway of haloperoxidases is not yet known, but it is widely accepted that the electrophilic species involved is the corresponding hypohalous acid (OH-X). **Scheme 6** presents the general mechanism for these enzymes; the halogen ion is oxidised by hydrogen peroxide, generating the highly reactive hypohalous acid intermediate, which is attacked by the organic nucleophile to produce the halogenated metabolite.<sup>110,112</sup>



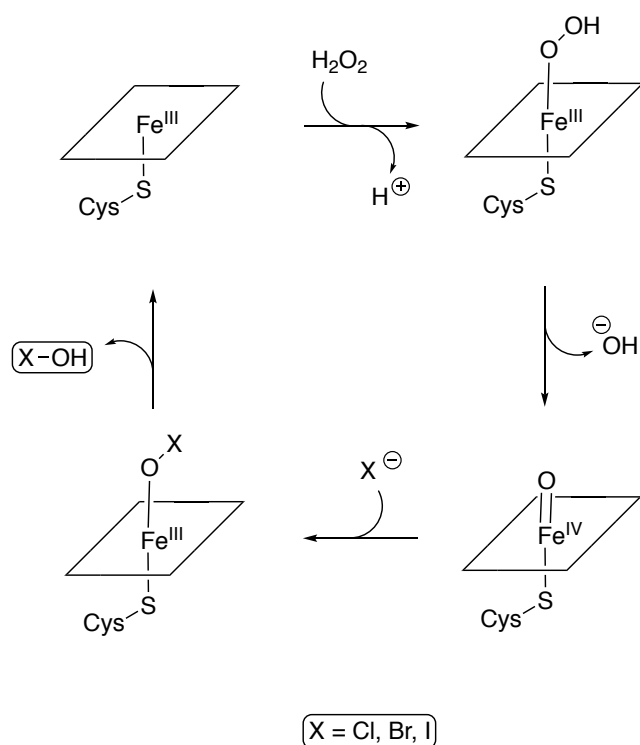
**Scheme 6.** General mechanism of haloperoxidases

The highly reactive hypohalous intermediate is most probably the cause of the low regiospecificity exhibited by haloperoxidases, which often generate a mixture of mono-, di- and trihalogenated metabolites.<sup>112</sup>

These peroxide-dependent enzymes are further divided in two categories: haem-iron and vanadium peroxidases. The mechanism for the formation of the electrophilic

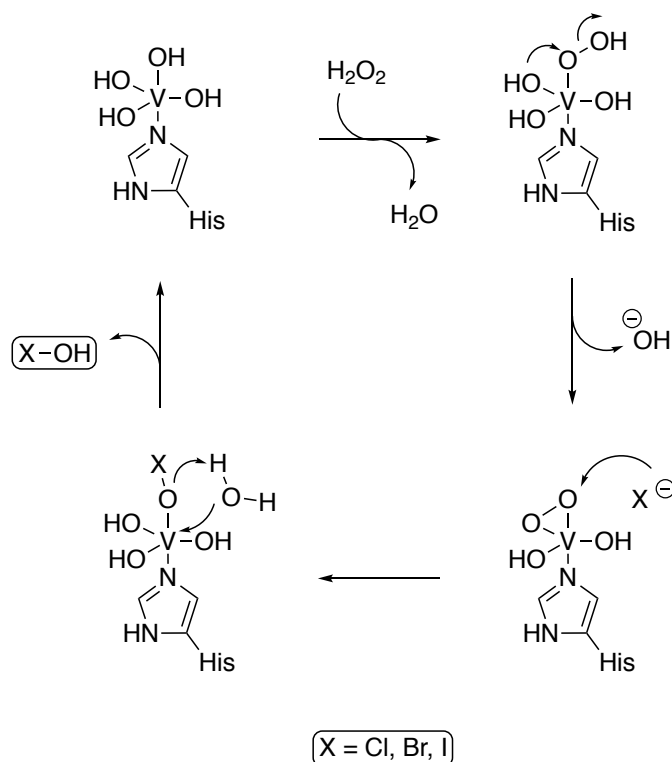
hypohalous acid in these sub-categories is not precisely known, but the hypothetical routes are shown in **Scheme 7** and **Scheme 8**.

In haem-iron haloperoxidases, the mechanistic route starts with the addition of hydrogen peroxide to  $\text{Fe}^{\text{III}}$ ; this complex is subsequently oxidised to the  $\text{Fe}^{\text{IV}}$  species, prior to the addition of the corresponding halide, re-establishing  $\text{Fe}^{\text{III}}$  and liberating hypohalous acid (**Scheme 7**).



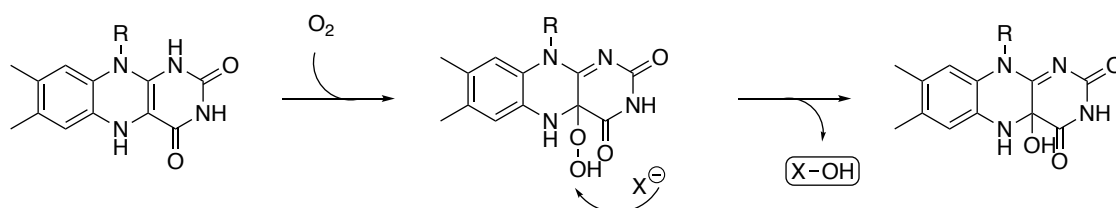
**Scheme 7.** Hypohalous acid formation by haem-iron haloperoxidases<sup>112</sup>

The route is similar for vanadium haloperoxidases. In that case, hydrogen peroxide binds to the metallic complex, before halide addition and final release of the hypohalous acid (**Scheme 8**).



**Scheme 8.** Hypohalous acid formation in vanadium-dependent haloperoxidases<sup>112</sup>

The exact mechanism by which flavin-dependent halogenases generate hypohalous acid may vary among the different members of the family. However, it is expected that the electrophilic agent is formed by the reaction of reduced flavin and molecular oxygen (**Scheme 9**).<sup>113, 112</sup>

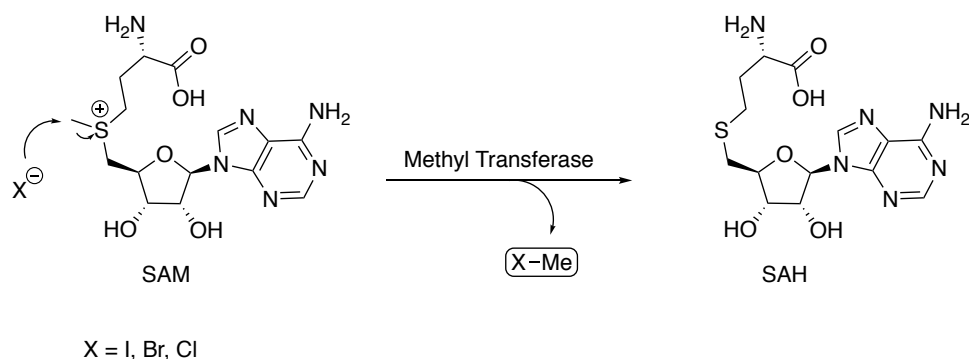


**Scheme 9.** Formation of the hypohalous acid intermediate in flavin-dependent halogenases

In 2005, Naismith and co-workers reported the existence of a 10 Å tunnel which separates the flavin from the substrate binding site, and through which the electrophilic hypohalous acid needs to migrate. This separates the generation of the electrophilic agent lysine residue from the actual site of the halogenation.<sup>114,115,110</sup> At the end of this tunnel is the  $\epsilon$ -NH<sub>2</sub> of a lysine, which has been implicated in the catalysis possibly *via* an intermediate R-NH<sub>2</sub>Cl species.

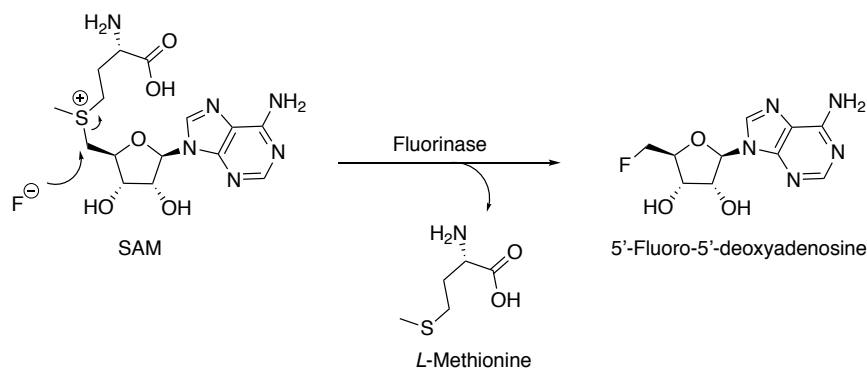
Nucleophilic halogenation is less common than the electrophilic route. All of the known enzymes involved in nucleophilic halogenation require *S*-adenosyl methionine (SAM) as a substrate.<sup>112</sup>

In the first case, the halide ( $X = \text{Cl}, \text{Br}, \text{I}$ ) attacks the sulfonium methyl group in SAM, in a methyltransfer reaction. The resulting product is the gaseous halomethane, as illustrated in **Scheme 10**.<sup>112</sup>



**Scheme 10.** Halomethanes production by the action of a methyl transferase

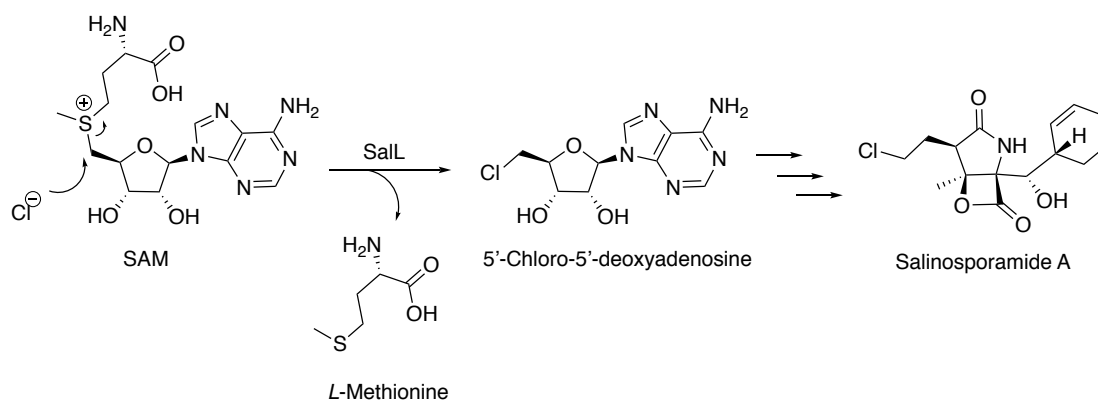
Nucleophilic halogenation can also involve the attack of the halide ion to the C5' position of SAM, displacing methionine and introducing the halogen.<sup>112</sup> The first example of this type of enzyme was reported in 2002 by O'Hagan *et al.*, isolated from *Streptomyces cattleya*.<sup>116</sup> This fascinating C-F bond-forming enzyme is named fluorodeoxyadenosine synthetase, although it is generally referred to as 'the fluorinase'.<sup>116</sup> **Scheme 11** shows the fluorinase-catalysed  $\text{S}_{\text{N}}2$ -type attack by fluoride ion to the C5' position of SAM, to generate 5'-fluoro-5'-deoxyadenosine (5'-FDA).<sup>117</sup> The relevant metabolic pathway will be discussed in **Section 1.5**.



**Scheme 11.** C-F bond formation catalysed by the fluorinase enzyme



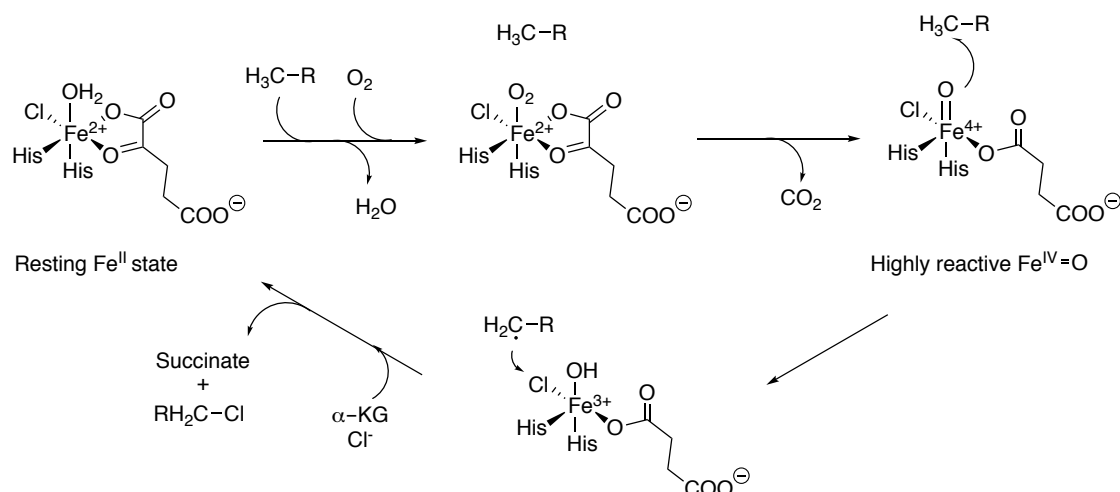
In 2009, Moore *et al.* reported a chlorinase analogous to the fluorinase.<sup>118</sup> This nucleophilic halogenating enzyme was found in the marine organism *Salinispora tropica*, and its metabolic pathway leads to the formation of salinosporamide A, a potent anticancer agent.<sup>118</sup> The first step of the metabolic pathway is presented in **Scheme 12**. More details of this pathway, and its relation to the fluorinase are described in **Section 1.5**.



**Scheme 12.** First step to the chlorinase pathway, and Salinosporamide A

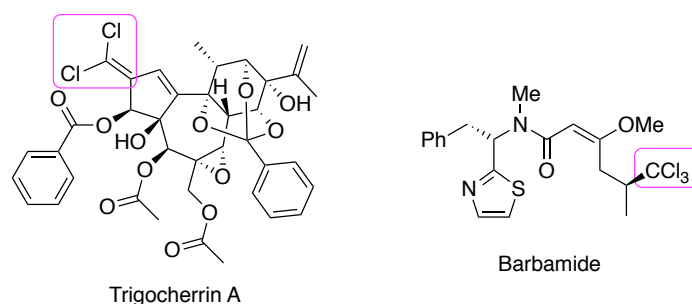
The third class of halogenating enzymes involves halogenation by single electron transfer. Given that these enzymes require the presence of  $\alpha$ -ketoglutarate as a co-substrate, they are often called  $\alpha$ -ketoglutarate-dependent halogenases; although they are also referred to as non-haem  $\text{Fe}^{\text{II}}$ - or  $\text{O}_2$ -dependent halogenases.<sup>115</sup>

Their mechanism was first investigated in the context of syringomycin E biosynthesis in *Pseudomonas syringae*.<sup>119, 120</sup> The proposed mechanism starts with an iron complex in its  $\text{Fe}^{\text{II}}$  resting state. The water molecule from this original complex is substituted by molecular oxygen, which is then reduced, oxidising the iron to  $\text{Fe}^{\text{IV}}$ . This newly formed species is highly reactive and attacks the substrate, generating a radical intermediate. The substrate radical is able to sequester the corresponding halide from the complex. These mechanistic steps are shown in **Scheme 13**.<sup>119</sup>



**Scheme 13.** Proposed mechanism for the  $\alpha$ -ketoglutarate-dependent halogenases<sup>119</sup>

An important characteristic of non-haem  $\text{Fe}^{\text{II}}$ -dependent halogenases is that they can halogenate unactivated, aliphatic carbon centres, sometimes leading to the formation of unusual halogenated motifs, such as vinylhalides.<sup>121</sup> Additionally, these halogenases can catalyse the transfer of multiple halogens within one active site.<sup>122</sup> The combination of these two properties results in the formation of natural products featuring several halogens on the same carbon atom; such is the case of trigocherrin A – which contains a dichlorovinyl moiety – or barbamide – containing a trichloromethyl group (**Fig. 17**).<sup>121</sup>



**Fig. 17** Unusual halogenated motifs generated by radical halogenation

### 1.4.3 Fluorinated natural products

Fluorine is the 13<sup>th</sup> most abundant element in the Earth's crust, and it appears primarily in minerals like fluorite, apatite and cryolite. However, the concentration of fluoride ion in seawater is extremely low. **Table 4** shows the occurrence of the different halogens in both seawater and the lithosphere.<sup>123</sup>

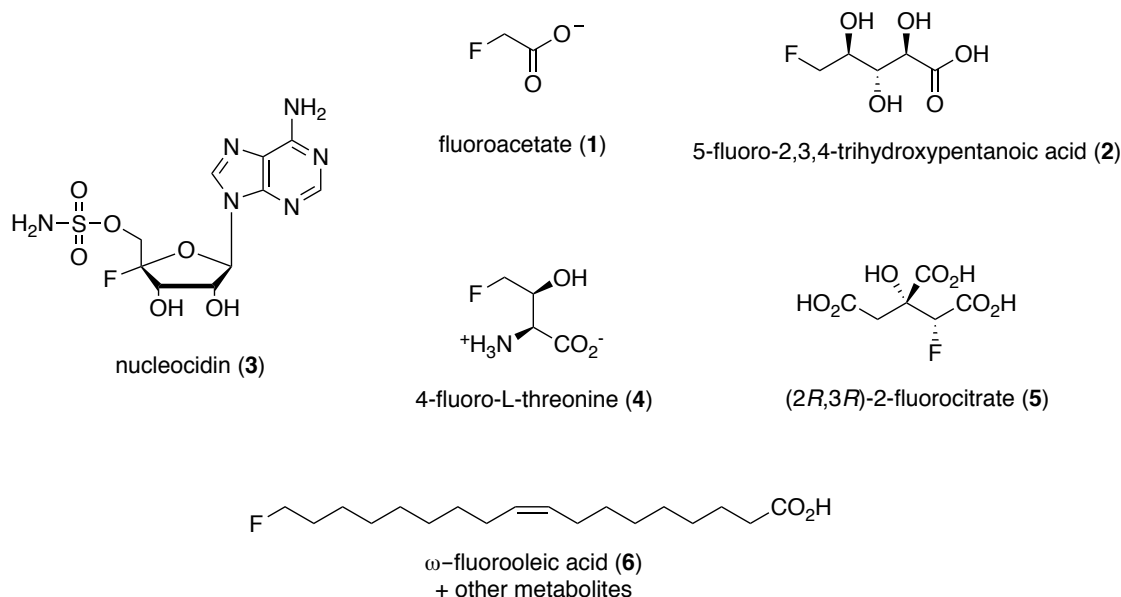
**Table 4.** Occurrence of the halogens on Earth

Element	Seawater (ppm)	Litosphere (ppm)
Fluorine	1.3	770
Chlorine	19000	550
Bromine	65	1.6
Iodine	0.05	0.3

In conjunction with low concentration in seawater, fluoride has the highest heat of hydration amongst the halogens (approximately 120 kcal/mol), and it follows that it is difficult to remove water molecules surrounding the anion and thus it is a poor nucleophile. Additionally, fluoride presents an extremely low oxidation potential (-2.87 eV), and it cannot be oxidised by haloperoxidases.<sup>124</sup>

These properties make fluorine – or fluoride – extremely difficult to enter into biochemistry and be incorporated by living organisms such as plants or bacteria. The first fluorinated natural product to be identified was fluoroacetate, in 1943.<sup>125</sup> Since then, only another five fluorine-containing structures from nature have been identified and isolated. In total, six types of fluorine-bearing natural products have been established, although the existence of some others is known although they have not been structurally characterised. Some reports of natural fluoro compounds have been discarded as unreliable.<sup>125</sup>

Fluoroacetate, (2*R*,3*R*)-2-fluorocitrate, 4-fluorothreonine, a range of  $\omega$ -fluorofatty acids, 5-fluoro-2,3,4-trihydroxypentanoic acid and nucleocidin have been correctly identified.



**Fig. 18** The six types of fluorinated natural products known to date

Fluoroacetate **1**, fluorocitrate **5** and the fluorofatty acids were all isolated from plants from Africa, and their discovery was closely related both in time and in content (research on fluoroacetate led to the discovery of fluorocitrate and the fluorofatty acids). The microbial products **2**, **3**, **4** and **6** came later.

Fluoroacetate was identified in 1943 by Marais in South Africa. The discovery was triggered by the investigations on *Dichapetalum cymosum*, a highly toxic indigenous plant that when consumed led to the death of foraging cattle. The local name of *D. cymosum* is 'gifblaar', which means 'poison leaf'.<sup>126</sup> Those investigations led to the isolation of fluoroacetate, the first fluorinated natural product, which was also the cause of the toxicity of gifblaar.

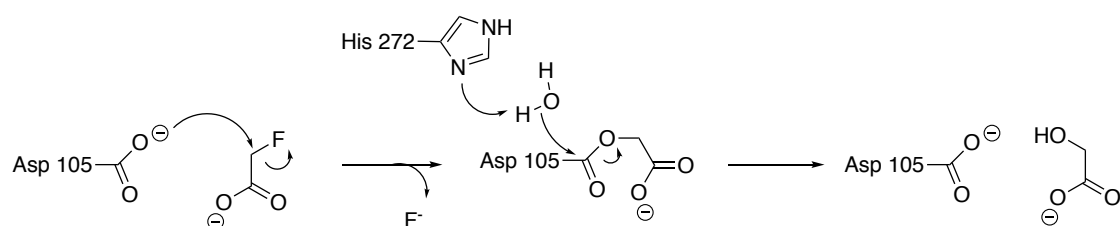
Fluoroacetate is also present in a variety of plants and microorganisms. As a curiosity, the highest level of fluoroacetate ever recorded was found in the seeds of the plant *Dichapetalum braunii*, which contained up to 8000 ppm.<sup>127,128</sup>

Nevertheless, several animals that live and forage in areas where fluoroacetate-producing plants are found, show resistance. An example of this behaviour has been observed for *Sindrus albimaculatus*, a caterpillar of the moth, which is capable of accumulating fluoroacetate. Their survival is due to an evolved resistance to the fluorinated toxin.<sup>129,130</sup>

In Australia, where this phenomenon has been extensively studied, one of the most striking cases is that of the emu. This seed-eating bird displays high resistance to fluoroacetate, with an LD<sub>50</sub> of up to 200 mg per kg,<sup>130</sup> whereas seed-eating birds grown and fed in areas where fluoroacetate is not naturally-occurring have an LD<sub>50</sub> of up to 20 mg per kg, one order of magnitude lower.<sup>130</sup>

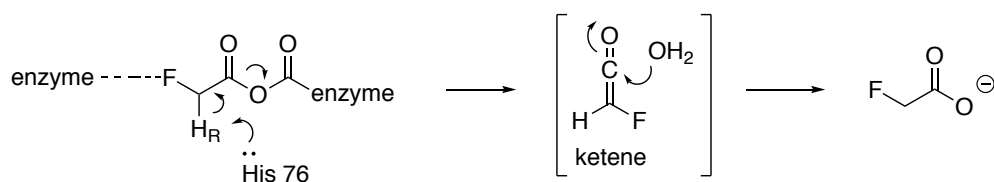
Given the complexity of these organisms, the exact mechanism by which they are able to cope with fluoroacetate is not known.<sup>131</sup>

Kawasaki and co-workers reported a defluorination enzymatic mechanism *Moraxella* species, shown in **Scheme 14**. The active site of the defluorinase contains an aspartate residue (Asp-105). The carboxylic group of Asp-105 acts as a nucleophile and displaces fluoride, forming an enzyme-bound intermediate. Then, His-272 activates water, and the resulting hydroxide ion hydrolyses the enzyme-bound intermediate, releasing glycolate.<sup>132</sup>



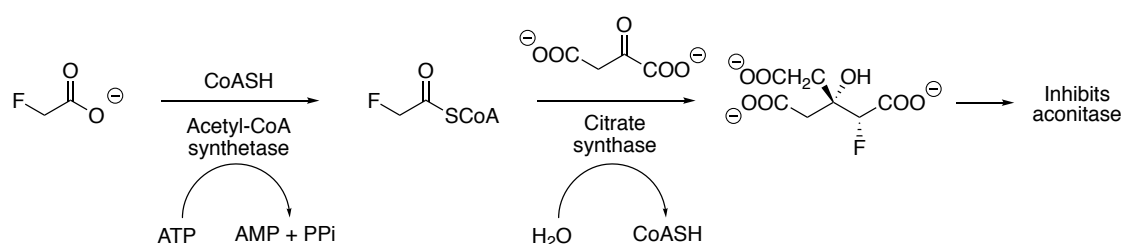
**Scheme 14.** Postulated enzymatic mechanism for *Moraxella* sp.

More recently, an investigation by Spencer discovered another bacterial fluoroacetate resistance enzyme in *Streptomyces cattleya*. In this case, it was a fluoroacetyl-CoA thioesterase (FIK).<sup>133</sup> Chang at Berkeley was able to shed some light on how this enzyme works by converting fluoroacetyl-CoA back to fluoroacetate, through a ketene intermediate (**Scheme 15**), and therefore avoiding the entrance of the toxin into the TCA cycle.<sup>134</sup>



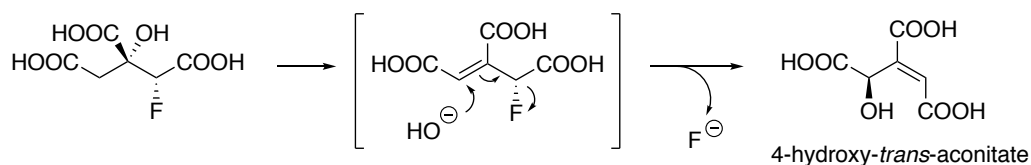
**Scheme 15.** Fluoroacetyl-CoA thioesterase's ketene intermediate

The toxicity of fluoroacetate is due to its *in vivo* conversion to (2*R*,3*R*)-2-fluorocitrate,<sup>135</sup> which is an inhibitor of aconitase<sup>136</sup> – a key enzyme of the citric acid cycle. **Scheme 16** shows the biotransformation of fluoroacetate into (2*R*,3*R*)-2-fluorocitrate, a procedure known as ‘lethal synthesis’.<sup>137</sup> The metabolic pathway starts with the action of an acetyl-CoA synthetase, which in the presence of ATP, converts fluoroacetate into fluoroacetyl-CoA. Fluoroacetyl-CoA is an active metabolite for the citric acid cycle, and is condensed with oxaloacetate by the action of a citrate synthase to yield 2-fluorocitrate, exclusively as the (2*R*,3*R*) stereoisomer – which is the only toxic diastereoisomer of the four possible isomers of 2-fluorocitrate.<sup>137,138,139</sup>



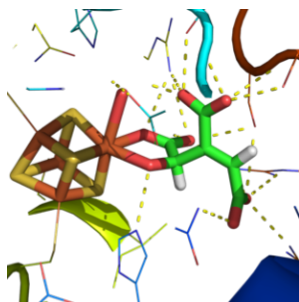
**Scheme 16.** Biotransformation of fluoroacetate into (2*R*,3*R*)-2-fluorocitrate

(2*R*,3*R*)-2-Fluorocitrate is a substrate for aconitase, which catalyses its transformation into isocitrate, through a *cis*-aconitate intermediate. This intermediate is attacked by hydroxide ion while bound to the enzyme, leading to the formation of 4-hydroxy-*trans*-aconitate, the actual inhibitor of aconitase (**Scheme 17**).<sup>140</sup>



**Scheme 17.** Transformation of (2*R*,3*R*)-2-fluorocitrate through the *cis*-aconitate intermediate into the *trans*-aconitate aconitase inhibitor

In 1996, Lauble and co-workers<sup>141</sup> solved the structure of the enzyme-bound inhibitor generated in this ‘lethal synthesis’ (4-hydroxy-*trans*-aconitate), and this image is reproduced in **Fig. 19**.



**Fig. 19** 4-Hydroxy-*trans*-aconitate bound to the enzyme with a [4Fe-4S] cluster

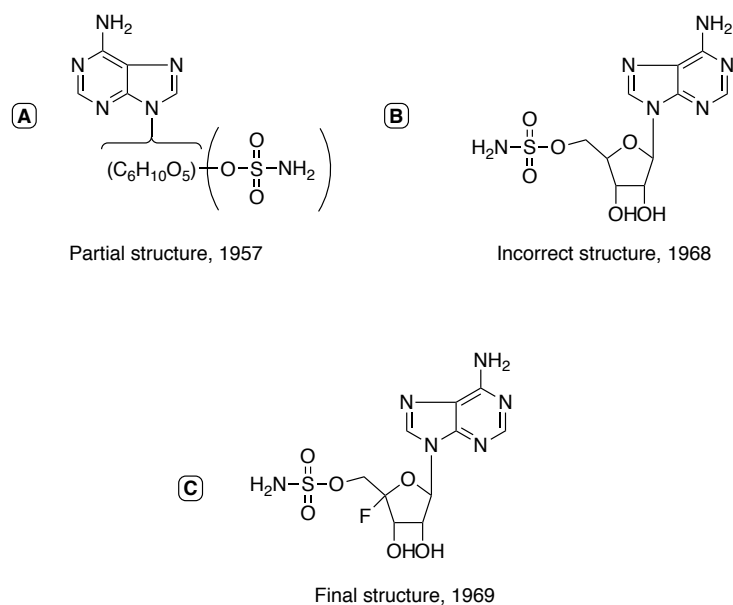
In 1959, Peters reported a toxic lipid from the seeds of *Dichapetalum toxicarium*, suggesting that it probably was a long chain  $\omega$ -fluorinated fatty acid.<sup>142</sup>

*D. toxicarium* is an indigenous plant to Sierra Leone, where the local people refer to it as ratsbane (rat poison), a very appropriate name, given that it is also used as a rodenticide. When the seeds of this plant are ingested, they cause vomiting, diarrhoea, and paralysis of the lower limbs, amongst other ailments. The paralysis of the lower extremities had afflicted the natives for generations, and was the cause for the plant's alternative name: 'broke back'.<sup>143</sup>

It was subsequently shown that the main component of the oil (source of 80% of the total organic fluorine present) was  $\omega$ -fluorooleic acid,<sup>144,145</sup> which co-occurred with at least two other components in the oil, as well as traces of fluoroacetate.<sup>146</sup>

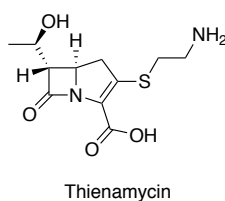
Nucleocidin was first isolated from *Streptomyces calvus* in 1957.<sup>147</sup> Both the elucidation of the structure and the biosynthesis of nucleocidin were hindered for many years, due to public strains being unable to produce the antibiotic.

The elucidation of nucleocidin's structure was not trivial. When initially isolated from *Streptomyces calvus*, a partial structure including an adenine moiety and an ester sulfamic acid was proposed (**Fig. 20a**).<sup>147</sup> A decade later, in 1968, the preliminary structure was corrected to 9-(4-*O*-sulfamoylpentofuranosyl)adenine (**Fig. 20b**),<sup>148</sup> and finally, in 1969, Morton *et al.* suggested the presence of fluorine in the C4' position of the ribose,<sup>149</sup> with a  $\beta$ -D-configuration.<sup>150</sup> The final structure of 4'-fluoro-5'-sulfamoyladosine (**Fig. 20c**) was confirmed by total synthesis in 1976.<sup>151</sup>



**Fig. 20** Different structures for nucleocidin, as proposed through its structural elucidation history

The discovery of new fluorinated natural products remained stalled for many years. In 1986, more than 25 years after Peters' identification of fluorofatty acids, Sanada *et al.* reported the isolation of 4-fluorothreonine and fluoroacetate, while attempting to improve the production of the  $\beta$ -lactam antibiotic thienamycin (**Fig. 21**), in *Streptomyces cattleya*.<sup>152</sup> Sanada discovered that when cultured in the presence of fluoride, *S. cattleya* was able to produce fluoroacetate, along with a novel fluorinated structure, which they identified as 4-fluorothreonine. This new fluorinated aminoacid exhibited antimicrobial activity.<sup>152</sup>



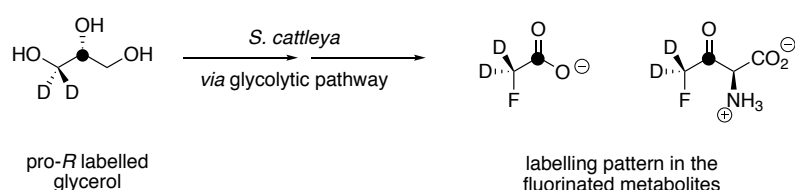
**Fig. 21** Antibiotic thienamycin

In an effort to locate the source of fluoroacetate and fluorothreonine from *S. cattleya*, Tamura and his colleagues designed a range of  $^{14}C$  and  $^{13}C$  labelling experiments.<sup>153</sup> While most of the envisaged experiments did not show successful results, the incubation of  $[2-^{13}C]$ -glycerol to the mycelium suspension caused enrichment of the carboxylic carbon of fluoroacetate with  $^{13}C$ . Further experiments with  $^{14}C$ -



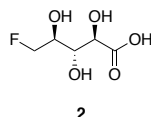
hydroxypyruvates showed that C-2 and C-3 are exclusively converted to fluoroacetate. These observations suggest that the carbon skeleton from fluoroacetate derives from glycerol through  $\beta$ -hydroxypyruvate.<sup>153</sup>

Separately, Nieschalk and O'Hagan studied the incorporation of deuterium-labelled glycerol into the fluorometabolites of *Streptomyces cattleya*. They used enantiotopically labelled (*R*)- and (*S*)-[1-<sup>2</sup>H<sub>2</sub>]-glycerol, and discovered that only the deuterium atoms from the pro-*R* side of glycerol were incorporated into fluoroacetate and fluorothreonine (**Scheme 18**).<sup>154</sup>



**Scheme 18.** Deuterium-labelled glycerol experiments in *S. cattleya*<sup>154</sup>

Almost thirty years passed before the isolation of a new fluorinated metabolite with the identification of 5-fluoro-2,3,4-trihydroxypentanoic acid **2** from *Streptomyces* sp. MA37. The structure of this compound was confirmed by total synthesis, and is shown in **Fig. 22**.<sup>155</sup>

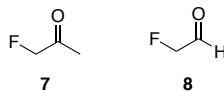


**Fig. 22** (2*R*,3*S*,4*S*)-5-Fluoro-2,3,4-trihydroxypentanoic acid from *Streptomyces* sp. MA37

Since the discovery of fluoroacetate in 1943, there have been several reports of new fluoroorganic molecules that were later confirmed as either wrongly analysed or incorrectly identified as natural products.

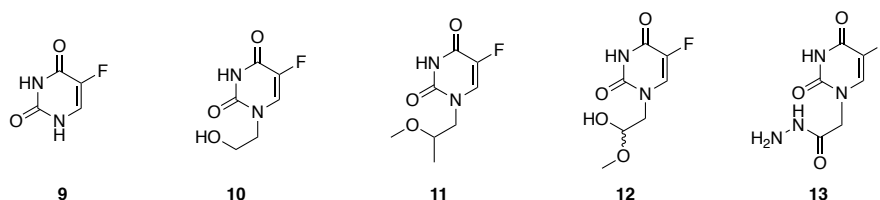
Peters and Shorthouse, reported in 1967 the observation of a ‘volatile form’ of a fluorometabolite, while investigating the metabolism of fluoroacetate in *Acacia georginae*.<sup>156</sup> By 1971, these scientists suggested that the ‘volatile form’ could be a fluoroketone. Since monofluoroacetone had previously been found in the livers of rats given fluoroacetate,<sup>157</sup> they concluded that the mysterious structure from *A. georginae* could be the same fluoroacetone **7**.<sup>158</sup> The authors were clear in expressing their doubts

on the identity of the ‘volatile form’. A widely accepted hypothesis is that the product they isolated was fluoroacetaldehyde **8**, an intermediate in the fluoroacetate pathway in *Streptomyces cattleya*.<sup>159</sup>



**Fig. 23** Fluoroacetone versus fluoroacetaldehyde

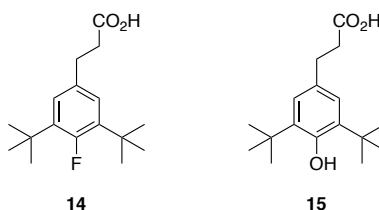
A more controversial case was presented with the discovery of some 5-fluorouracil derivatives from the sponge *Phakellia fusca* in the South Sea of China.<sup>160</sup> In 2003, Xu *et al.* isolated the five 5-fluorouracil derivatives displayed in **Fig. 24**.



**Fig. 24** 5-Fluorouracil derivatives isolated from the South Sea of China

Given that derivatives **9** and **10** are well-known anti-tumour agents,<sup>161</sup> it is most likely that these isolates are just the result of industrial contamination in the ocean, absorbed by the sponge.<sup>162,163</sup>

More recently, the isolation of di-*tert*-butylfluorophenylpropionic acid **14** from *Streptomyces* sp. TC1 was reported.<sup>164</sup> Despite the claim that the molecule contained a fluorine atom, no spectral evidence was provided, apart from a very broad <sup>19</sup>F NMR signal in the Supporting Information of the article.<sup>164</sup> O’Hagan *et al.* synthesised a sample of **14**, and comparison with the published spectra proved the molecules to be different.<sup>165</sup> Instead, they propose that the original isolate was phenolpropionic acid **15**, a well-known antioxidant.<sup>165</sup> Both structures are shown in **Fig. 25**.



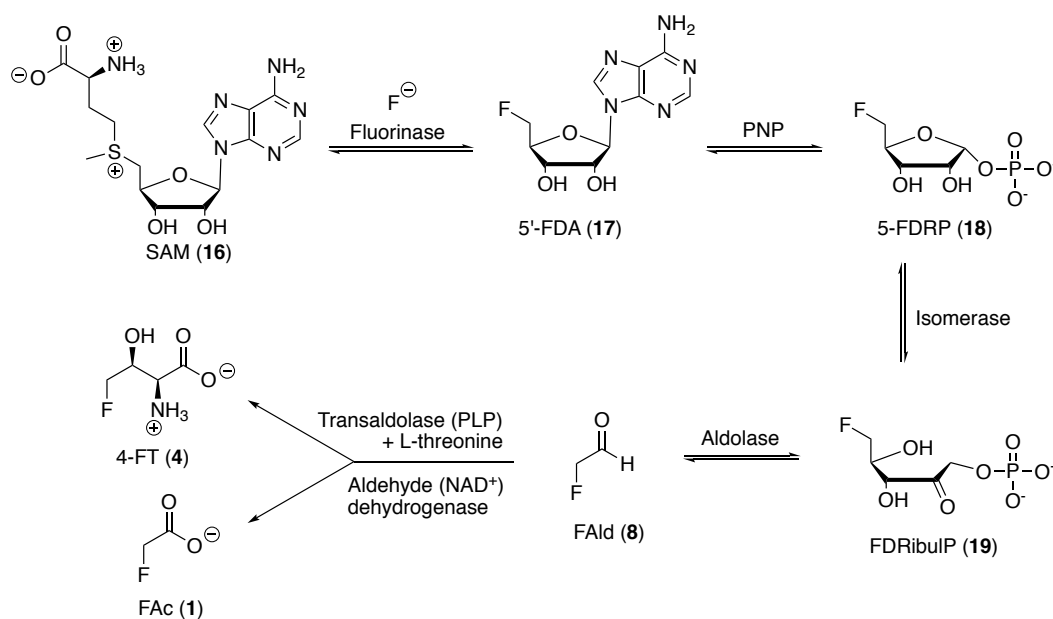
**Fig. 25** Di-*tert*-butylfluorophenylpropionic acid and the phenolic derivative

## 1.5 The fluorinase

### 1.5.1 Discovery and crystal structure of the first fluorinating enzyme

The natural sources of organofluorine products have been extensively studied, both because of fluoroacetate's toxicity and their rare occurrence. In 2002, O'Hagan *et al.* identified the first fluorinating enzyme, from the bacterium *Streptomyces cattleya* that produces 4-fluorothreonine.<sup>116</sup> The enzyme is able to catalyse the formation of a carbon-fluorine bond by the  $S_N2$ -like attack of fluoride to the C5' position of SAM, to generate 5'-FDA.<sup>116</sup> This intermediate is transformed into fluoroacetaldehyde, which is oxidised to fluoroacetate.<sup>116</sup> While the production of fluoroacetate and 4-fluorothreonine by *Streptomyces cattleya* was originally observed by Sanada, no relationship could be established at the time on the formation of these products.<sup>152</sup>

The biosynthetic pathway to fluoroacetate and 4-fluorothreonine in *S. cattleya* involves five enzymatic steps, and a branch point, which separates the formation of fluoroacetate and 4-fluorothreonine.<sup>166</sup> The reactions are shown in **Scheme 19**.

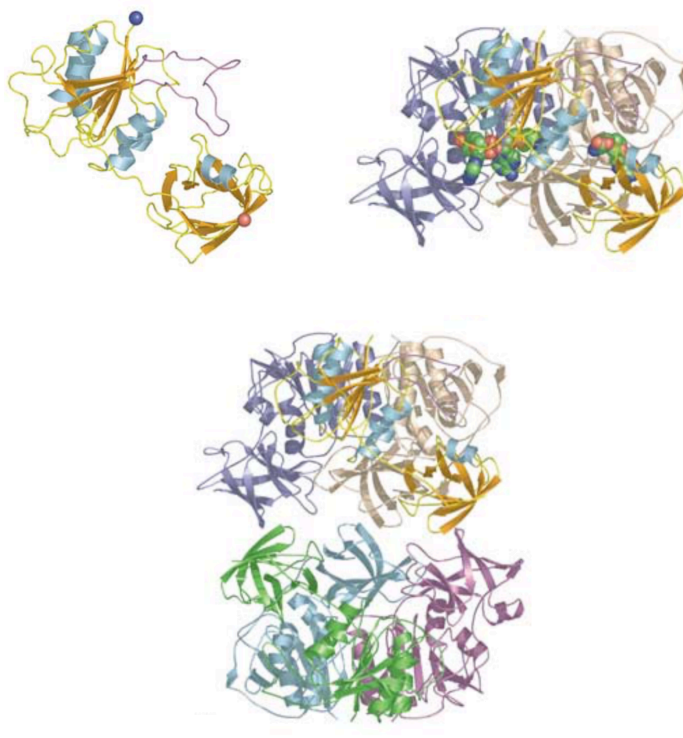


**Scheme 19.** Fluorinase pathway in *Streptomyces cattleya*

The pathway shown in **Scheme 19** starts with the  $S_N2$  attack of fluoride on SAM, catalysed by the fluorinase, to yield to 5'-FDA. 5'-FDA is then transformed into 5-fluorodeoxyribose phosphate (5-FDRP) by the action of a purine nucleoside

phosphorilase (PNP). An isomerase then converts 5-FDRP into 5-fluorodeoxyribulose phosphate (FDRibulP), which is the substrate for the aldolase that generates fluoroacetaldehyde **8**. Fluoroacetaldehyde is a branch point on the metabolic pathway. A pyridoxal phosphate (PLP) dependent transaldolase, yields 4-*L*-fluorothreonine in combination with *L*-threonine. Alternatively, an NAD<sup>+</sup>-dependent aldehyde dehydrogenase transforms fluoroacetaldehyde into fluoroacetate **1**.<sup>166</sup>

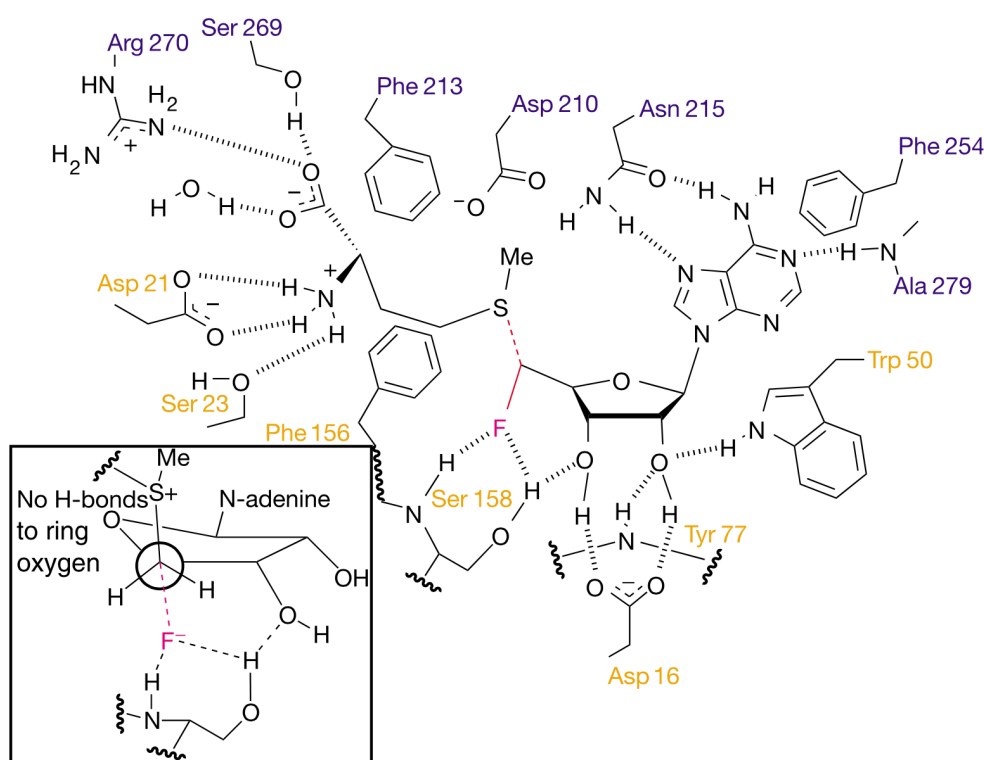
Purification and *in vitro* over-expression of 5'-fluoro-5'-deoxyadenosine synthase allowed a deeper understanding of the enzyme's kinetics, and opened the possibility of reproducing the protein's sequence by PCR, in order to determine the complete coding sequence for the *flA* gene.<sup>117</sup> Crystallography was used to determine the structure of the wild-type enzyme, isolated from *S. cattleya*.<sup>117</sup> Crystal structures showed that the fluorinase is constructed as a dimer of trimers – forming a hexamer. Each monomer unit is organised by the amino-terminal domain, linked to a smaller carboxy-terminal domain; and by a 'distinctive, extended loop of 21 amino acid residues'.<sup>117</sup> **Fig. 26** shows the structure of the monomeric units, the trimers and the hexamers determined by X-ray crystallography.<sup>117</sup>



**Fig. 26** Crystal structures of the monomeric units, trimers and hexamers of the fluorinase<sup>117</sup>

Structures with a bound SAM molecule show that the substrate binds to the interface between the C-terminal and the N-terminal domains, and the trimer holds three bound molecules.<sup>117</sup>

Isotopic labelling experiments which resulted in prochiral deuterium labelling at C-5' of SAM, incubated in whole-cell experiments, indicated an 'inversion of configuration' during the  $F^-$  attack on the C5' position of SAM.<sup>167</sup> The hypothesis was consistent with the crystal structure, showing 5'-FDA and methionine bound to the enzyme, and displaying an *anti* relationship between the C-F bond (in red in **Fig. 27**) and the C-S bond (dotted red, **Fig. 27**).<sup>117</sup>



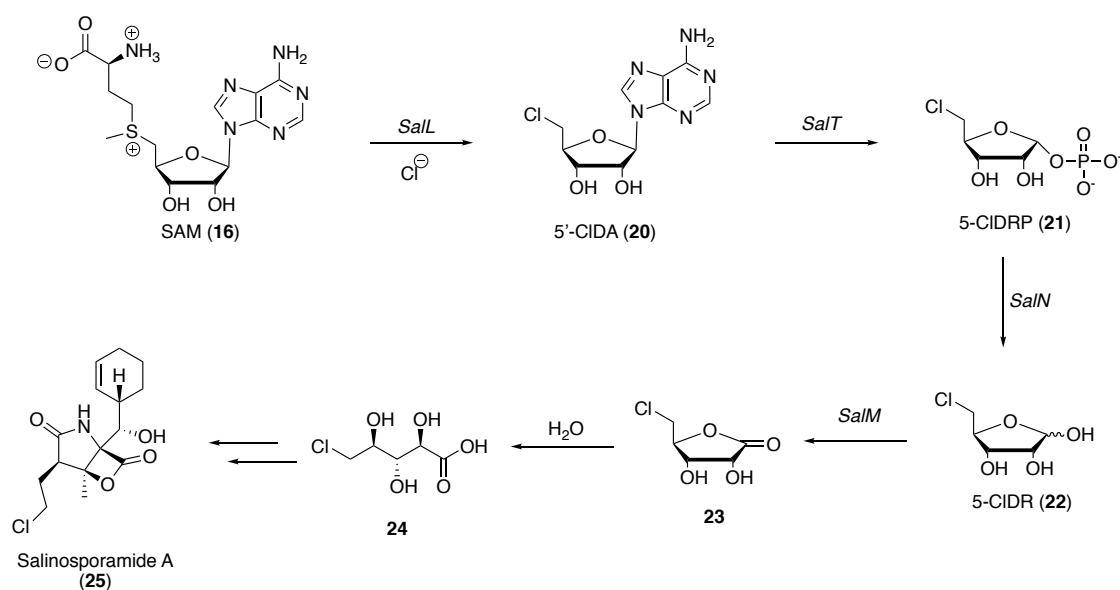
**Fig. 27** Crystal structure of 5'-FDA formation, in a  $S_N2$  fashion<sup>117</sup>

### 1.5.3 *Streptomyces* sp. MA37

In 2014, a fluorinase gene homologue was identified in a genome isolated from *Streptomyces* sp. MA37 found in Ghana. It showed an 87% analogy to the original fluorinase.<sup>168</sup> In the presence of fluoride ion, the cultures of MA37 also produced

fluoroacetate and 4-fluorothreonine, as well as a range of unidentified fluorinated products, as observed by  $^{19}\text{F}$ -NMR spectroscopy.<sup>168</sup>

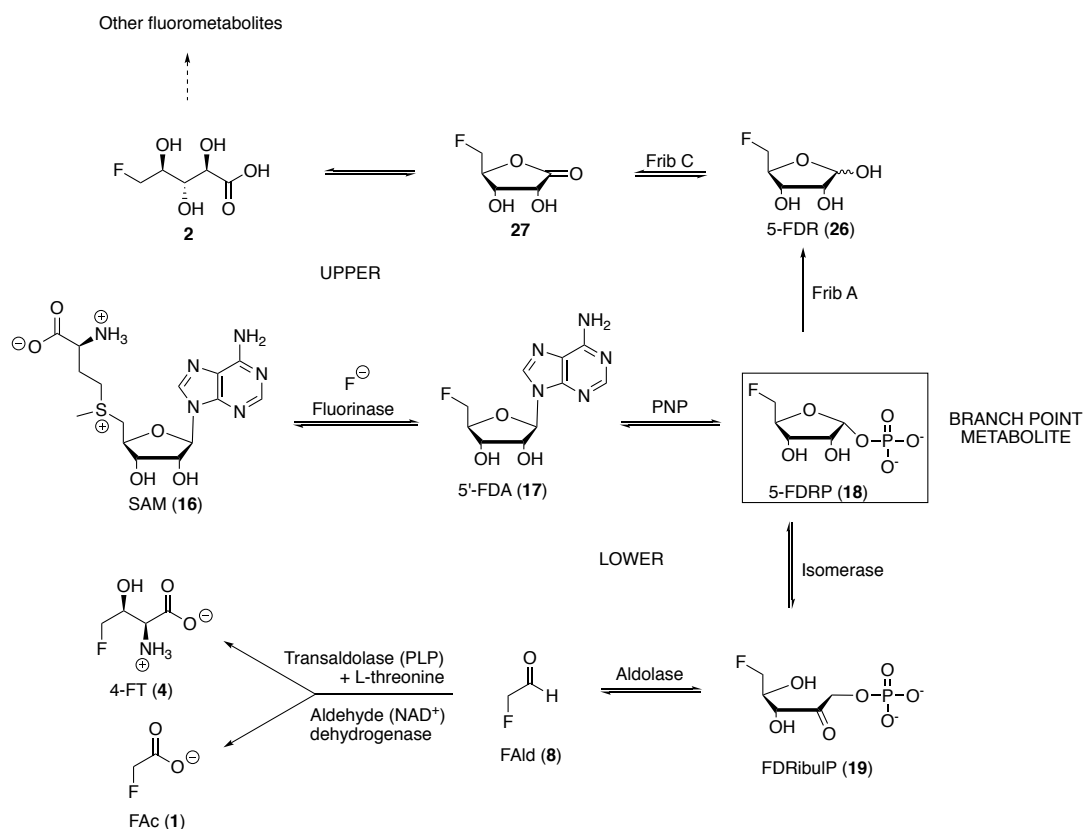
Further investigations showed that the genome in MA37 had a high homology to that of *Salinispora tropica*, the organism with a chlorinase enzyme involved in the biosynthesis of salinosporamide A.<sup>118</sup> This chlorinase is closely related to the fluorinase, and they work in a similar manner. The early steps of the salinosporamide A production are shown in **Scheme 20**.<sup>118</sup>



**Scheme 20.** Early steps in the salinosporamide A chlorinase metabolic pathway<sup>118</sup>

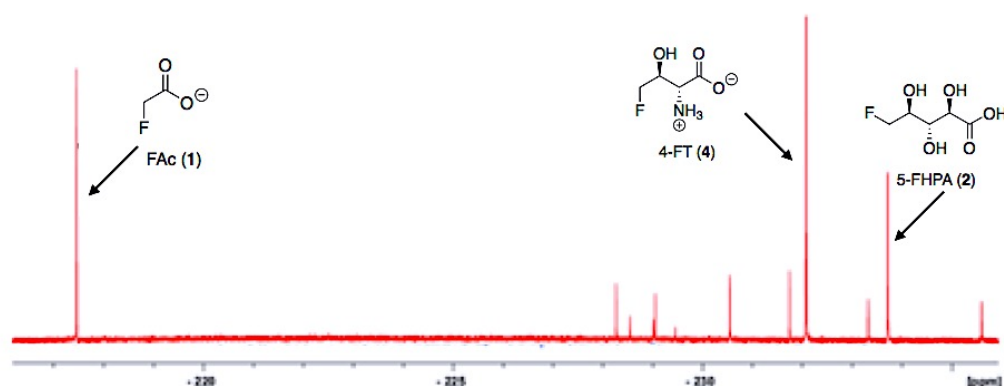
Unlike *Streptomyces cattleya* and *Streptomyces* sp. MA37, *Salinispora tropica* does not produce fluoroacetate or fluorothreonine. However, the first two steps of the metabolic pathway are analogous to those in *S. cattleya*. Comparison of the genome of *S. tropica* and MA37 showed functional homology amongst some of the enzymes involved in both metabolic pathways.<sup>118,155</sup>

Therefore, a metabolic pathway for the new fluorometabolites was proposed, merging those of *S. cattleya* and *S. tropica*, and creating a branch point in 5-FDRP (18).<sup>155</sup> This is shown in **Scheme 21**, with the separated upper and lower branches signalled.<sup>155</sup>



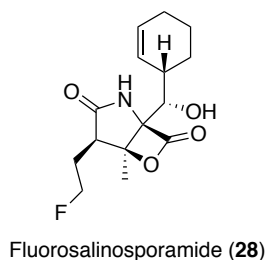
**Scheme 21.** Proposed metabolic pathway for the fluorinase *Streptomyces* sp. MA37<sup>155</sup>

Testing this hypothesis led to the discovery of the most-recently isolated fluorometabolite: (2*R*,3*S*,4*S*)-5-fluoro-2,3,4-trihydroxypentanoic acid (**2**). The identity of this new compound was proven by total synthesis and GC-MS analysis. So far, the pathway is structurally secure up to the formation of **2**, but the nature of the further metabolites remains unknown.<sup>155</sup>



**Fig. 28** <sup>19</sup>F NMR analysis of a *Streptomyces* sp. MA37 culture, with identified molecules **1**, **2** and **4** and several unidentified structures<sup>155</sup>

It is interesting to note that although the fluorinase can act as a chlorinase,<sup>169</sup> the opposite does not occur, and the chlorinase is not capable of incorporating fluorine. Genetic engineering - carried out in a collaboration between St Andrews and the Scripps Institution of Oceanography - have enabled the formation of fluorosalinosporamide, in an engineered *Salinispora tropica*, by knocking out that chlorinase gene and replacing it with the fluorinase gene.<sup>170</sup>



**Fig. 29** Structure of fluorosalinosporamide

#### 1.5.4 Other organisms containing a fluorinase

Since the discovery of the fluorinase in *Streptomyces* sp. MA37, another three fluorinases have been identified by genome mining. All present more than 80% homology to the original fluorinase in *Streptomyces cattleya*.<sup>168,171</sup>

Two of the new fluorinases were found in the genomes of sequenced organisms (*Actinoplanes* sp. N902-109 and *Nocardia brasiliensis*). PCR amplification of the genes, as well as over expression and purification of the enzymes in *E. coli* proved that both fluorinases are functional.<sup>168</sup>

Additionally, a third functional fluorinase was found in *Streptomyces xinghaiensis*, the first from a marine source. The fluorinase pathway in *S. xinghaiensis* leads to the formation of fluoroacetate, but not fluorothreonine, and the production of fluoroacetate is sea-salt dependent.<sup>171</sup>



## 1.6 References

---

- <sup>1</sup> a) H. Moissan, 'Action d'un courant électrique sur l'acide fluorhydrique anhydre', *Comptes rendus hebdomadaires des séances de l'Académie des Sciences*, **1886**, 102, 1543. b) H. Moissan 'Sur la décomposition de l'acide fluorhydrique par un courant électrique', *Comptes rendus hebdomadaires des séances de l'Académie des Sciences*, **1886**, 103, 202. c) H. Moissan, 'Recherches sur l'isolement du fluor', Gauthier-Villars, **1887**.
- <sup>2</sup> A. E. Tutton, *Nature*, **1887**, 37, 179.
- <sup>3</sup> E. Frémy, *Ann. Chim. Phys.* **1856**, 3, 5.
- <sup>4</sup> The Nobel Prize in Chemistry 1906. NobelPrize.org. Nobel Media AB 2018. <<https://www.nobelprize.org/prizes/chemistry/1906/summary/>>
- <sup>5</sup> a) A. M. Ampère; *Ann. Chim. Phys.* **1816**, 2, 5. b) J. Joubert (A. M. Ampère); *Ann. Chim. Phys.* **1885**, 4, 5. c) W. H. Waggoner; *J. Chem. Educ.* **1976**, 53, 27.
- <sup>6</sup> G. Agricola, *De Re Metallica*. Translated from the first latin edition (**1556**) by H. C. Hooper and L. H. Cooper (**1950**), Dover Publications Inc. New York.
- <sup>7</sup> F. van Catz Smalenburg, *Leerboek der scheikunde* (Volume 1), page 230. A. En J. Honkoop, 1827. (Unofficially translated by David van Brussel – University of St Andrews)
- <sup>8</sup> H. Davy, *Phil. Trans. Royal Soc.* **1813**, 103, 263.
- <sup>9</sup> B. G. Aston, J. A. Harrell, I. Shaw, 'Ancient Egyptian Materials and Technology', Ed. Paul T. Nicholson, Ian Shaw, Cambridge University Press, **2006** (4th edition), page 31.
- <sup>10</sup> A. E. Werner, M. Bimson; Some opacifying agents in Oriental Glass, in *Advances in glass technology*, part 2. Eds: F. R. Matson, G. E. Rindone), Plenum Press, New York, **1963**, 303.
- <sup>11</sup> J. Henderson, J. An, H. Ma; *Archaeometry*, **2018**, 60, 88.

- 
- <sup>12</sup> R. McLaughlin, 'The Roman Empire and the Silk Routes: The Ancient World Economy and the Empires of Parthia, Central Asia and Han China'. Pen and Sword, 2016, page 169.
- <sup>13</sup> K. Kannaiyan, M. Jyothi, S. B. Gujjar, K. Vasudha, J. Murgesh, S. A. Pai; *Int. J. Sci. Res.* **2018**, 7, 63.
- <sup>14</sup> L. R. McDowell, 'Mineral Nutrition History: The Early Years' (**2017**), First Edition Desing Publishing, Inc., page 554.
- <sup>15</sup> A. S. Marggraf, *Mém. Acad. Berlin*, **1768**, 3.
- <sup>16</sup> C. W. Scheele, *Mém. Acad. Stockholm*, **1771**, 33, 120.
- <sup>17</sup> J. Wisniak, *Indian J. Chem. Technol.* **2002**, 9, 363.
- <sup>18</sup> L. J. Thénard, J. L. Gay-Lussac; *Ann. Chim.* **1808**, 68, 169.
- <sup>19</sup> P. Louyet, *Compt. Rendv.* **1846**, 23, 960.
- <sup>20</sup> J. Emsley, 'Nature's Building Blocks: An A-Z Guide to the Elements'. Oxford University Press, **2003**, p 147.
- <sup>21</sup> H. Davy, *Philos. Trans.* **1813**, 103, 263.
- <sup>22</sup> G. Gore, *Philosophical Transactions of the Royal Society of London*, **1869**, 159, 173.
- <sup>23</sup> A. Harsanyi, G. Sandford; *Green Chem.* **2015**, 17, 2081.
- <sup>24</sup> The Holy Bible. Exodus 10:13-15.
- <sup>25</sup> a) M. L. McKinney, R. M. Schoch, L. Yonavjak; 'Environmental Science: Systems and Solutions', Jones and Bartlett Learning, **2007**. b) J. Unsworth, History of Pesticide Use, agrochemicals.iupac.org, **2010**.
- <sup>26</sup> M. F. Raisbeck, Chapter 66 'Organochlorine Pesticides', Small Animal Toxicology (3rd Edition), **2013**, p 709.
- <sup>27</sup> V. Turusov, V. Rakitsky, L. Tomatis; *Environ. Health Perspect.* **2002**, 110, 125.

- 
- <sup>28</sup> R. Jayaraj, P. Megha, P. Sreedev; *Interdiscip. Toxicol.* **2016**, 9, 90.
- <sup>29</sup> P. Jeschke, *Pest Manag. Sci.* **2010**, 66, 10.
- <sup>30</sup> Fluorine and the Environment. Agrochemicals, Archaeology, Green Chemistry and Water. Edited by Alan Tressaud, Elsevier, 2006. Chapter 4. Fluorine-containing agrochemicals: an overview of recent developments. G. Theodoridis.
- <sup>31</sup> D. W. Grainger, C. W. Stewart; *Chapter 1. Fluorinated coatings and films*. Fluorinated surfaces, coatings and films. Castner and Grainger, ACS Symposium Series, Washington DC, 2001.
- <sup>32</sup> [www.ptfecoatings.com/technology/what-is-teflon.php](http://www.ptfecoatings.com/technology/what-is-teflon.php)
- <sup>33</sup> [www.solvay.com/en/products/brands/fomblin-pfpe-lubricants](http://www.solvay.com/en/products/brands/fomblin-pfpe-lubricants)
- <sup>34</sup> P. Kirsch; *J. Fluor. Chem.* **2015**, 177, 29.
- <sup>35</sup> M. Hird; *Chem. Soc. Rev.* **2007**, 36, 2070.
- <sup>36</sup> a) N. Al-Maharik, P. Kirsch, A. M. Z. Slawin, D. B. Cordes, D. O'Hagan; *Org. Biomol. Chem.* **2016**, 14, 9974. b) S. Yamada, S. Hashishita, T. Asai, T. Ishihara, T. Konno; *Org. Biomol. Chem.* **2017**, 15, 1495. c) M. Spengler, R. Y. Dong, C. A. Michal, M. Pfletscher, M. Giese; *J. Mater. Chem. C*, **2017**, 5, 2235.
- <sup>37</sup> J. J. Vaquero, P. Kinahan; *Annu. Rev. Biomed. Eng.* **2015**, 17, 385.
- <sup>38</sup> K. Duncan; *J. Nucl. Med. Technol.* **1998**, 26, 228.
- <sup>39</sup> E. Miele, G. P. Spinelli, F. Tomao, A. Zullo, F. De Marinis, G. Pasciuti, L. Rossi, F. Zoratto, S. Tomao; *J. Exp. Clin. Cancer Res.* **2008**, 27, 52.
- <sup>40</sup> S. Dall'Angelo, N. Bandaranayaka, A. D. Windhorst, D. J. Vugts, D. Van der Born, M. Onega, L. F. Schweiger, M. Zanda, D. O'Hagan; *Nuc. Med. Biol.* **2013**, 40, 464.
- <sup>41</sup> K. Müller, C. Faeh, F. Diederich; *Science*, **2007**, 317, 1881.
- <sup>42</sup> Y. Zhou, J. Wang, Z. Gu, S. Wang, W. Zhu, J. L. Aceña, V. A. Soloshonok, K. Izawa, H. Liu; *Chem. Rev.* **2016**, 116, 422.

- 
- <sup>43</sup> a) J. Fried, E. F. Sabo; *J. Am. Chem. Soc.* **1954**, 76, 1455. b) D. R. Friend, G. W. Chang; *J. Med. Chem.* **1985**, 28, 51.
- <sup>44</sup> K. Schultebrasucks, K. Wingenfeld, C. Otte, M. Quinkler; *Neuroendocrinology*, **2016**, 103, 315.
- <sup>45</sup> S. Purser, P. R. Moore, S. Swallow, V. Gouverneur; *Chem. Soc. Rev.* **2008**, 37, 320.
- <sup>46</sup> G. T. Schumock, E. C. Li, K. J. Suda, M. D. Wiest, J. Stubbings, L. M. Matusiak, R. J. Hunkler, L. C. Vermeulen; *Am. J. Health Syst. Pharm.* **2016**, 73, 1058.
- <sup>47</sup> D. O'Hagan; *J. Fluor. Chem.* **2010**, 131, 1071.
- <sup>48</sup> J. Wang, M. Sánchez-Roselló, J. L. Aceña, C. Del Pozo, A. E. Sorochinsky, S. Fustero, V. A. Soloshonok, H. Liu; *Chem Rev.* **2014**, 114, 2432.
- <sup>49</sup> H.-J. Böhm, D. Banner, S. Bendels, M. Kansy, B. Kuhn, K. Müller, U. Obst-Sander, M. Stahl; *ChemBioChem*, **2004**, 5, 637.
- <sup>50</sup> M. van Heek, C. F. France, D. S. Compton, R. L. McLeod, N. P. Yumibe, K. B. Alton, E. J. Sybertz, H. R. Davis Jr.; *J. Pharmacol. Exper. Ther.* **1997**, 283, 157.
- <sup>51</sup> S. B. Rosenblum, T. Huynh, A. Afonso, H. R. Davis Jr., N. Yumibe, J. W. Clader, D. A. Burnett; *J. Med. Chem.* **1998**, 41, 973.
- <sup>52</sup> M. E. Sweeney, R. R. Johnson; *Expert Opin. Drug Met.* **2007**, 3, 441.
- <sup>53</sup> S. Swallow, Chapter 5: Fluorine-containing pharmaceuticals. V. Gouverneur, K. Müller; *Fluorine in Pharmaceutical and Medicinal Chemistry*, Molecular Medicine and Medicinal Chemistry, Volume 6, Imperial College Press, London, **2012**.
- <sup>54</sup> B. E. Smart; *J. Fluor. Chem.* **2001**, 109, 3.
- <sup>55</sup> N. Muller; *J. Pharm. Sci.* **1986**, 75, 987.
- <sup>56</sup> B. E. Smart; *J. Fluorine Chem.* **2001**, 109, 3.
- <sup>57</sup> R. T. Jacobs, P. R. Bernstein, L. A. Cronk, E. P. Vacek, L. F. Newcomb, D. Aharony, C. K. Buckner, E. J. Kusner; *J. Med. Chem.* **1994**, 37, 1282.

- 
- <sup>58</sup> J.-P. Bégué, D. Bonnet-Delpon; Chapter 13. ‘Biological Impacts of Fluorination: Pharmaceuticals Based on Natural Products’, p. 565. A. Tressaud, G. Haufe; Fluorine and Heath, Elsevier, **2008**.
- <sup>59</sup> Q. A. Huchet, B. Kuhn, B. Wagner, H. Fisher, M. Kansy, D. Zimmerli, E. M. Carreira, K. Müller; *J. Fluorine Chem.* **2013**, *152*, 119.
- <sup>60</sup> a) C. A. Lipinski, F. Lombardo, B. W. Dominy, P. J. Feeney; *Adv. Drug Deliv. Rev.* **1997**, *23*, 3. b) C. A. Lipinski, *Adv. Drug Deliv. Rev.* **2016**, *101*, 34.
- <sup>61</sup> C. Vraha, L. Nies, K.-H. Wagner, M. Hacker, W. Wadsak, M. Mitterhauser; *NucMedBio*, **2017**, *50*, 1.
- <sup>62</sup> H. Pajouhesh, G. R. Lenz; *NeuroRx*. **2005**, *2*, 541.
- <sup>63</sup> C.-T. Lu, Y.-Z. Zhao, H. L. Wong, J. Cai, L. Peng, X.-Q. Tian; *Int. J. Nanomedicine*, **2014**, *9*, 2241.
- <sup>64</sup> C. Hansch, A. R. Steward, S. M. Anderson, D. Bentley; *J. Med. Chem.* **1967**, *11*, 1.
- <sup>65</sup> S. K. Bhal, ‘Log *P* – Making sense of the Value’, application note, Advanced Chemistry Development Inc., Toronto, Canada.
- <sup>66</sup> A. B. Khan, T. Kingsley, P. Caroline; *J. Pharm. Res.* **2017**, *16*, 257.
- <sup>67</sup> K. R. Vora, W. I. Higuchi, N. F. Ho; *Int. J. Pharm. Sci. Invent.* **1972**, *61*, 1785.
- <sup>68</sup> E. P. Gillis, K. J. Eastman, M. D. Hill, D. J. Donnelly, N. A. Meanwell; *J. Med. Chem.* **2015**, *58*, 8315.
- <sup>69</sup> J. A. DiMasi, H. G. Grabowski, R. W. Hansen; *J. Health Econ.* **2016**, *47*, 20-33.
- <sup>70</sup> Z. Zhang, W. Tang; *Acta Pharmaceutica Sinica B*, **2018**, *8*, 721-732.
- <sup>71</sup> R. T. Williams; *Gut*, **1972**, *13*, 579.
- <sup>72</sup> T. T. T. Nguyen, Y.-J. Choi, H. B. Lee; *Mycobiology*, **2017**, *45*, 318.
- <sup>73</sup> S. Asha, M. Vidyavathi; *Biotechnol. Adv.* **2009**, *27*, 16.

- 
- <sup>74</sup> R. O. Brennan, B. J. Crain, A. M. Proctor, D. T. Durack; *Am. J. Clin. Pathol.* **1983**, *80*, 98.
- <sup>75</sup> R. A. Samson; *Proc. K. Ned. Akad. Wet. C*, **1969**, *72*, 322.
- <sup>76</sup> U. Baijal, B. S. Mehrotra; *Sydowia*, **1980**, *33*, 1.
- <sup>77</sup> J. A. Lunn, W. A. Shipton; *Trans. Br. Mycol. Soc.* **1983**, *81*, 303.
- <sup>78</sup> I. Weitzman, M. Y. Crist; *Mycologia*, **1979**, *71*, 1024.
- <sup>79</sup> I. Weitzman; *Trans. Br. Mycol. Soc.* **1984**, *83*, 527.
- <sup>80</sup> K. C. Marshall, M. Alexander; *J. Bacteriol.* **1960**, *80*, 412.
- <sup>81</sup> M. T. Green, J. H. Dawson, H. B. Gray; *Science*, **2004**, *304*, 1653.
- <sup>82</sup> J. T. Groves; *Nature Chemistry*, **2014**, *6*, 89.
- <sup>83</sup> D. H. Peterson, H. C. Murray; *J. Am. Chem. Soc.* **1952**, *74*, 1871.
- <sup>84</sup> R. Beukers, A. F. Marx, M. H. J. Zuidweg; ‘Microbial conversion as a tool in the preparation of drugs’, Editor: E. J. Ariens, ‘Drug Design, vol. III’. New York Academic Press, 1972.
- <sup>85</sup> a) R. E. Betts, D. E. Walters, J. P. Rosazza; *J. Med. Chem.* **1974**, *17*, 599. b) R. V. Smith, J. P. Rosazza; *J. Pharm. Sci.* **1975**, *64*, 1737. c) R. V. Smith, J. P. Rosazza; *J. Nat. Prod.* **1983**, *46*, 79.
- <sup>86</sup> Y. W. J. Wong, P. J. Davis; *Pharm. Res.* **1989**, *6*, 982.
- <sup>87</sup> J. D. Moody, D. Zhang, T. M. Heinze, C. E. Cerniglia; *Appl. Environ. Microbiol.* **2000**, *66*, 3646.
- <sup>88</sup> C.-J. Cha, D. R. Doerge, C. E. Cerniglia; *Appl. Environ. Microbiol.* **2001**, *67*, 4358.
- <sup>89</sup> J. Amadio, C. D. Murphy; *Biotechnol. Lett.* **2010**, *33*, 321.
- <sup>90</sup> J. Amadio, K. Gordon, C. D. Murphy; *Appl. Environ. Microbiol.* **2010**, *76*, 6299.
- <sup>91</sup> J. Amadio, C. D. Murphy; *Appl. Microbiol. Biotechnol.* **2010**, *86*, 345.

- 
- <sup>92</sup> J. Amadio, E. Casey, C. D. Murphy; *Appl. Microbiol. Biotechnol.* **2013**, *97*, 5955.
- <sup>93</sup> T. V. Bright, F. Dalton, V. L. Elder, C. D. Murphy, N. K. O'Connor, G. Sandford; *Org. Biomol. Chem.* **2013**, *11*, 1135.
- <sup>94</sup> E. Kavanagh, M. Winn, C. Nic Gabhann, N. K. O'Connor, P. Beier, C. D. Murphy; *Environ. Sci. Pollut. Res.* **2014**, *21*, 753.
- <sup>95</sup> L. Quinn, R. Dempsey, E. Casey, A. Kane, C. D. Murphy; *J. Ind. Microbiol. Biotechnol.* **2015**, *42*, 799.
- <sup>96</sup> A. S. Hampton, L. Mikulski, W. Palmer-Brown, C. D. Murphy, G. Sandford; *Bioorg. Med. Chem. Lett.* **2016**, *26*, 2255.
- <sup>97</sup> W. Palmer-Brown, B. Dunne, Y. Ortin, M. A. Fox, G. Sandford, C. D. Murphy; *Xenobiotica*, **2017**, *47*, 763.
- <sup>98</sup> E. Stone; *Phil. Trans.* **1763**, *53*, 195.
- <sup>99</sup> K. C. Nicolaou, E. J. Sorensen, N. Wissinger; *J. Chem. Educ.* **1998**, *75*, 1225.
- <sup>100</sup> A. Fleming, *Br. J. Exp. Pathol.* **1929**, *10*, 226.
- <sup>101</sup> The Nobel Prize in Physiology or Medicine 1945. NobelPrize.org. Nobel Media AB2018. <<https://www.nobelprize.org/prizes/medicine/1945/summary/>>
- <sup>102</sup> K. Yamada; Cobalt: Its Role in Health and Disease. In: Sigel A., Sigel H., Sigel R. (eds) *Interrelations between Essential Metal Ions and Human Diseases. Metal Ions in Life Sciences*, vol 13. Springer, Dordrecht, **2013**.
- <sup>103</sup> a) D. Crowfoot-Hodgkin, A. W. Johnson, A. R. Todd; *Chem. Soc. Spec. Publ.* **1955**, *3*, 109. b) D. Crowfoot-Hodgkin, J. Kamper, M. MacKay, J. Pickworth, K. N. Trueblood, J. G. White; *Nature*, **1956**, *178*, 64.
- <sup>104</sup> A. Eschenmoser; *Q. Rev.* **1970**, *24*, 366.
- <sup>105</sup> The Nobel Prize in Chemistry 1964. NobelPrize.org. Nobel Media AB 2018. <<https://www.nobelprize.org/prizes/chemistry/1964/summary/>>
- <sup>106</sup> E. Drechsel; *Zeitschrift fuer Biology*, **1896**, *33*, 85.

- 
- <sup>107</sup> G. W. Gribble; *J. Nat. Prod.* **1992**, *55*, 1353.
- <sup>108</sup> G. W. Gribble; *Naturally Occurring Organohalogen Compounds – A Comprehensive Update*, Wien: Springer, **2010**.
- <sup>109</sup> G. W. Gribble; *Mar. Drugs*; **2015**, *13*, 4044.
- <sup>110</sup> C. S. Neumann, D. G. Fujimori, C. T. Walsh; *Chem. Biol.* **2008**, *15*, 99.
- <sup>111</sup> L. P. Hager, P. D. Shaw; *J. Am. Chem. Soc.* **1959**, *81*, 1011.
- <sup>112</sup> D. R. M. Smith, S. Grüşchow, R. J. M. Goss; *Curr. Op. Chem. Biol.* **2013**, *17*, 276.
- <sup>113</sup> S. Flecks, E. P. Patallo, X. Zhu, A. J. Ernyei, G. Seifert, A. Schneider, C. Dong, J. H. Naismith, K.-H. van Pée; *Angew. Chem. Int. Ed.* **2008**, *47*, 9533.
- <sup>114</sup> C. Dong, S. Flecks, S. Unversucht, C. Haupt, K.-H. van Pée, J. H. Naismith; *Science*, **2005**, *309*, 2216.
- <sup>115</sup> C. S. Neumann, C. T. Walsh, R. R. Kay; *PNAS*, **2010**, *107*, 5798.
- <sup>116</sup> D. O'Hagan, C. Schaffrath, S. L. Cobb, J. T. G. Hamilton, C. D. Murphy; *Nature*, **2002**, *416*, 279.
- <sup>117</sup> C. Dong, F. Huang, H. Deng, C., J. B. Spencer, D. O'Hagan, J. H. Naismith; *Nature*, **2004**, *427*, 561.
- <sup>118</sup> A. S. Eustáquio, R. McGlinchey, Y. Liu, C. Hazzard, L. L. Beer, G. Florova, M. M. Alhamadsheh, A. Lechner, A. J. Kale, Y. Kobayashi, K. A. Reynolds, B. S. Moore; *PNAS*, **2009**, *106*, 12295.
- <sup>119</sup> F. H. Vaillancourt, J. Yin, C. T. Walsh; *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 10111.
- <sup>120</sup> a) C. Krebs, D. G. Fujimori, C. T. Walsh, J. M. Bollinger Jr. *Acc. Chem. Res.* **2007**, *40*, 484. b) C. Wagner, M. El Omari, G. M. König; *J. Nat. Prod.* **2009**, *72*, 540.
- <sup>121</sup> P.-M. Allard, M.-T. Martin, M.-E. Dau, P. Leyssen, F. Gueritte, M. Litaudon; *Org. Lett.* **2012**, *14*, 342.



- 
- <sup>122</sup> C. Seibold, H. Schnerr, J. Rumpf, A. Kunzendorf, C. Hatscher, T. Wage, A. J. Ernyei, C. J. Dong, J. H. Naismith, K.-H. van Pée; *Biocatal. Biotransform.* **2006**, *24*, 401.
- <sup>123</sup> K. L. Kirk; Chapter 1. The Halogens: Discovery, occurrence and biochemistry of the free elements. *Biochemistry of the Elements*, vol. 9A+B, **1991**, page 2.
- <sup>124</sup> D. O'Hagan, H. Deng; *Chem Rev.* **2015**, *115*, 634.
- <sup>125</sup> H. Deng, D. O'Hagan, C. Schaffrath; *Nat. Prod. Rep.* **2004**, *21*, 773.
- <sup>126</sup> J. S. C. Marais; *Onderstepoort J. Vet. Sci. Anim. Ind.* **1943**, *18*, 203.
- <sup>127</sup> D. O'Hagan, R. Perry, J. M. Lock, J. J. M. Meyer, L. Dasaradhi, J. T. G. Hamilton, D. B. Harper; *Phytochemistry*, **1993**, *33*, 1043.
- <sup>128</sup> D. B. Harper, D. O'Hagan; *Nat. Prod. Rep.* **1994**, *11*, 123.
- <sup>129</sup> J. J. M. Meyer, D. O'Hagan; *Chem. Br.* **1992**, *28*, 785.
- <sup>130</sup> D. O'Hagan, D. B. Harper; *J. Fluor. Chem.* **1999**, *100*, 127.
- <sup>131</sup> L. E. X. Leong, S. Khan, C. K. Davis, S. E. Denman, C. S. McSweeney; *J. Anim. Sci. Biotechnol.* **2017**, *8*, 55.
- <sup>132</sup> H. Kawasaki, N. Tore, K. Tonomura; *Agric. Biol. Chem.* **1981**, *45*, 29.
- <sup>133</sup> a) F. Huang, S. F. Haydock, D. Spiteller, T. Mironenko, T. L. Li, D. O'Hagan, P. F. Leadlay, J. B. Spencer; *Chem. Biol.* **2006**, *13*, 475. b) M. V. B. Dias, F. Huang, D. Y. Chirgadze, M. Tosin, D. Spiteller, E. F. V. Dry, P. F. Leadlay, J. B. Spencer, T. L. Blundell; *J. Biol. Chem.* **2010**, *285*, 22495.
- <sup>134</sup> A. M. Weeks, N. S. Keddie, R. D. P. Wadoux, D. O'Hagan, M. C. Y. Chang; *Biochemistry*, **2014**, *53*, 2053.
- <sup>135</sup> a) C. Liébecq, R. A. Peters; *Biochim. Biophys. Acta*, **1949**, *3*, 215. b) C. Martius; *Liebigs Ann. Chem.* **1949**, *561*, 227. c) P. Buffa, R. A. Peters, *J. Physiol. Lond.* **1950**, *110*, 488.
- <sup>136</sup> W. D. Lotspeich, R. A. Peters, T. H. Wilson; *Biochem. J.* **1951**, *48*, 467.

- 
- <sup>137</sup> R. A. Peters, *Proc. R. Soc. Lond. B*, **1952**, *139*, 143.
- <sup>138</sup> R. A. Peters, R. W. Wakelin, P. Buffa, L. C. Thomas; *Proc. R. Soc. Lond. B*, **1953**, *140*, 497.
- <sup>139</sup> W. B. Elliott, G. Kalnitsky; *J. Biol. Chem.* **1950**, *186*, 487.
- <sup>140</sup> W. D. Lotspeich, R. A. Peters, T. H. Wilson; *Biochem. J.* **1952**, *51*, 20.
- <sup>141</sup> H. Lauble, M. C. Kennedy, M. H. Emptage, H. Beinert, C. D. Stout; *Proc. Natl. Acad. Sci. USA*, **1996**, *93*, 699.
- <sup>142</sup> R. A. Peters, R. W. Wakelin, A. J. P. Martin, J. Webb, F. T. Birks; *Biochem. J.* **1959**, *71*, 245.
- <sup>143</sup> H. D. Neuwinger; 'African Ethnobotany: Poisons and Drugs: Chemistry, Pharmacology, Toxicology', **1996**, CRC Press, p. 415.
- <sup>144</sup> R. A. Peters, R. J. Hall; *Biochem. Pharmac.* **1959**, *2*, 25.
- <sup>145</sup> R. A. Peters, R. J. Hall, P. F. V. Ward; *Biochem. J.* **1960**, *77*, 17.
- <sup>146</sup> P. F. V. Ward, R. J. Hall, R. A. Peters; *Nature*, **1964**, *201*, 611.
- <sup>147</sup> C. W. Waller, J. B. Patrick, W. Fulmor, W. E. Meyer; *J. Am. Chem. Soc.* **1957**, *79*, 1011.
- <sup>148</sup> J. B. Patrick, W. E. Meyer; Abstracts, 156<sup>th</sup> National Meeting of the American Chemical Society, Atlantic City, N. J. **1968**.
- <sup>149</sup> G. O. Morton, J. E. Lancaster, G. E. Van Lear, W. Fulmor, W. E. Meyer; *J. Am. Chem. Soc.* **1969**, *91*, 1535.
- <sup>150</sup> D. A. Schuman, R. K. Robins, M. J. Robins; *J. Am. Chem. Soc.* **1969**, *91*, 3391.
- <sup>151</sup> I. D. Jenkins, J. P. H. Verheyden, J. G. Moffatt; *J. Am. Chem. Soc.* **1976**, *98*, 3346.
- <sup>152</sup> M. Sanada, T. Miyano, S. Iwadare, J. M. Williamson, B. H. Arison, J. L. Smith, A. W. Douglas, J. M. Liesch, E. Inamine; *Antibiotics*, **1986**, *39*, 259.
- <sup>153</sup> T. Tamura, M. Wada, N. Esaki, K. Sato; *J. Bacteriol.* **1995**, *177*, 2265.

- 
- <sup>154</sup> J. Neischalk, J. T. G. Hamilton, C. D. Murphy, D. B. Harper, D. O'Hagan; *Chem. Commun.* **1997**, 799.
- <sup>155</sup> L. Ma, A. Bartholomé, M. H. Tong, Z. Qin, T. Shepherd, K. Kyeremeh, H. Deng, D. O'Hagan; *Chem. Sci.* **2015**, 6, 1414.
- <sup>156</sup> R. A. Peters, M. Shorthouse; *Nature*, **1967**, 216, 80.
- <sup>157</sup> K. Miura, S. Otsuka, K. Honda; *Bull. Agric. Chem. Soc. Japan*, **1956**, 20, 219.
- <sup>158</sup> R. A. Peters, M. Shorthouse, *Nature*, **1971**, 231, 123.
- <sup>159</sup> S. J. Moss, C. D. Murphy, J. T. G. Hamilton, W. C. McRoberts, D. O'Hagan, C. Schaffrath, D. B. Harper; *Chem. Commun.* **2000**, 22, 2281.
- <sup>160</sup> X.-H. Xu, G.-M. Yao, Y.-M. Li, J.-H. Lu, C.-J. Lin, X. Wang, C.-H. Kong; *J. Nat. Prod.* **2003**, 66, 285.
- <sup>161</sup> S. Spiegelman, R. Sawyer, R. Nayak, E. Ritzi, R. Stolfi, D. Martin; *Proc. Natl. Acad. Sci. USA*, **1980**, 77, 4966.
- <sup>162</sup> S. Ozaki, *Med. Res. Rev.* **1996**, 16, 51.
- <sup>163</sup> D. A. Hopwood, *Natural Product Biosynthesis by Microorganisms and Plants*, Part 2, page 223.
- <sup>164</sup> N. Jaivel, C. Uvarani, R. Rajesh, D. Velmurugan, P. Marimuthu; *J. Nat. Prod.* **2014**, 77, 2.
- <sup>165</sup> M. S. Ayoup, D. B. Cordes, A. M. Z. Slawin, D. O'Hagan; *J. Nat. Prod.* **2014**, 77, 1249.
- <sup>166</sup> a) S. L. Cobb, H. Deng, J. T. G. Hamilton, R. P. McGlinchey, D. O'Hagan; *Chem. Commun.* **2004**, 0, 592. b) M. Onega, R. P. McGlinchey, H. Deng, J. T. G. Hamilton, D. O'Hagan; *Bioorg. Chem.* **2007**, 35, 375.
- <sup>167</sup> D. O'Hagan, R. J. M. Goss, A. Meddour, J. Courtieu; *J. Am. Chem. Soc.* **2003**, 125, 379.

- 
- <sup>168</sup> H. Deng, L. Ma, N. Bandaranayaka, Z. Qin, G. Mann, K. Kyeremeh, Y. Yu, T. Shepherd, J. H. Naismith, D. O'Hagan; *ChemBioChem*, **2014**, *15*, 364.
- <sup>169</sup> H. Deng, S. L. Cobb, A. R. McEwan, R. P. McGlinchey, J. H. Naismith, D. O'Hagan, D. A. Robinson, J. B. Spencer; *Angew. Chem. Int. Ed.* **2006**, *45*, 759.
- <sup>170</sup> A. S. Eustáquio, D. O'Hagan, B. S. Moore; *J. Nat. Prod.* **2010**, *73*, 378.
- <sup>171</sup> S. Huang, L. Ma, M. H. Tong, Y. Yu, D. O'Hagan, H. Deng; *Org. Biol. Chem.* **2014**, *12*, 4828.

## Chapter 2. Metabolism and lipophilicity profile of a range of novel fluorinated motifs

Organofluorine compounds are relevant across all fields of performance chemicals, including medicinal, materials and agricultural chemistry. The substitution of hydrogen atoms for fluorine analogues bestows new characteristics to the recipient molecule, often improving its pharmaco-kinetic properties.<sup>1,2,3</sup> The exploration of new fluorinated motifs is therefore of interest to offer new building blocks for research programmes.

In this context, lipophilicity is a key pharmaco-kinetic property, normally expressed as a Log *P* value. Lipophilicity is directly related to a drug's efficacy, since it determines the permeability of the compound through the cell membrane and its distribution in albumin.<sup>2</sup> Historically, medicinal chemists aim for Log *P* values lower than five, but over the years, and with the levels of exigency growing considerably, this value has been lowered to three.<sup>4</sup>

Many authors have studied and discussed the effect that fluorine has on the lipophilicity of a drug,<sup>1,5,6</sup> and while it is generally assumed that fluorination increases lipophilicity, this does not always hold.<sup>7</sup> The extent of that affirmation was explored for our new fluorinated motifs, and will be an important section of this chapter.

In drug design programmes there is an additional need to study and analyse how a candidate behaves within the human body, in order to assess how viable that particular compound would be in metabolic terms. The fungus *Cunninghamella elegans* has been used as a model for mammalian drug metabolism, and it emerges as a very useful tool, given its ease of evaluation.<sup>8</sup>

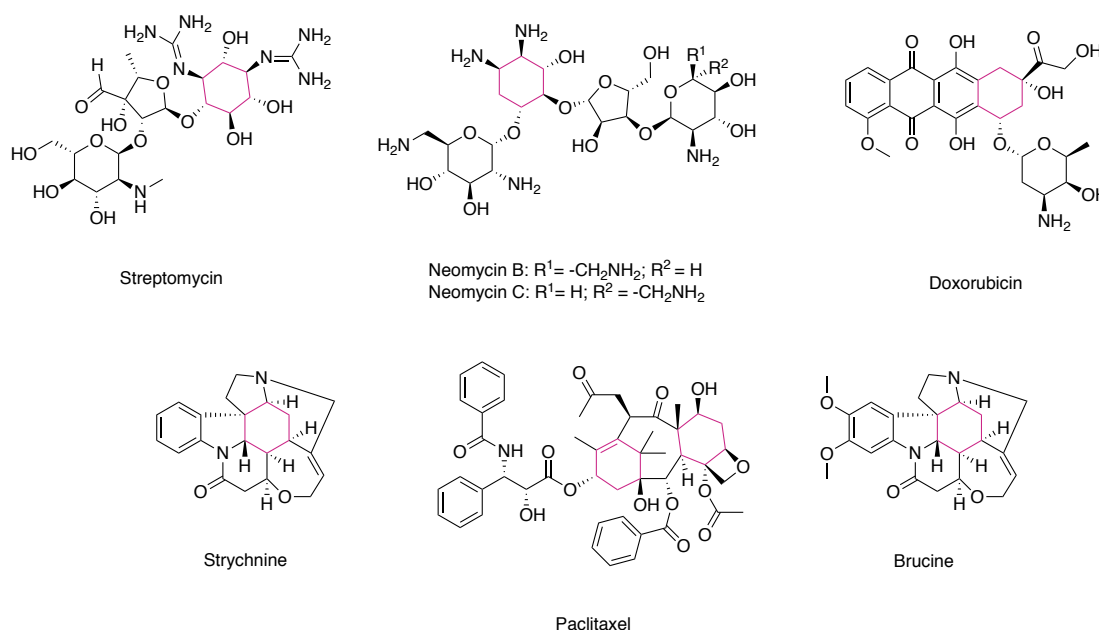
*Cunninghamella elegans* contains a high level of oxidation enzymes of the cytochrome P<sub>450</sub> class, enzymes which have a role in phase I drug metabolism.<sup>9,10</sup> Previous reports with this organism also describe the appearance of conjugative metabolites, resulting from phase II enzymes, for more labile compounds.<sup>11</sup>

During the last few years, different members from the St Andrews group have developed and synthesised a range of novel fluorinated motifs. Among the new motifs there are fluorocyclopropanes and fluorocyclohexanes, and also  $\alpha,\alpha$ -difluoroethyl thio- and oxy- ethers. The study of their metabolic fate, as well as research exploring their lipophilicity profile, is the main objective of this Chapter.

## 2.1 Derivatives of all-*cis* phenylfluorocyclohexane

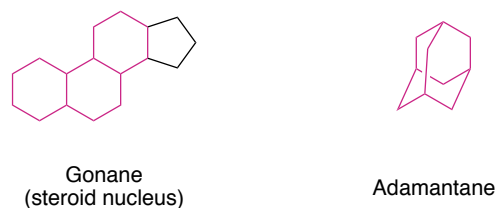
### 2.1.1 Introduction to the syntheses and potential applications of all-*cis* (fluorocyclohexyl)benzenes

Of all the cycloalkanes, cyclohexane is the most extensively studied as a motif in organic chemistry. Understanding the fundamentals of cyclohexane's conformation, by Barton<sup>12</sup> and Hassel,<sup>13</sup> was awarded the Nobel Prize for Chemistry in 1969.<sup>14</sup> The size of its ring makes it optimal in terms of strain and stability. Cyclohexane moieties are present in many natural products. The range goes from antibiotics, such as streptomycin (isolated from *Streptomyces – Actinomyces – griseus* in 1943)<sup>15</sup> and neomycin (firstly found in *Streptomyces fradiae – Actinomyces fradii* – in 1949)<sup>16</sup>, to poisons – like strychnine or its derivative brucine. Other natural products include chemotherapeutic drugs like doxorubicin or paclitaxel.



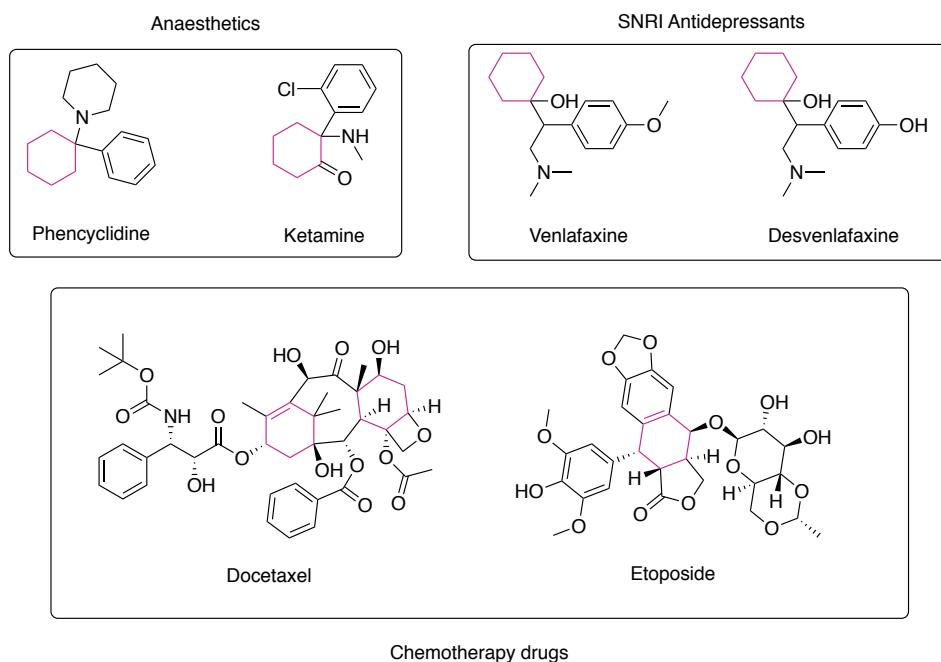
**Fig. 1** Examples of natural products featuring one cyclohexane ring

Fused cyclohexanes are a fundamental feature of steroids – whose nucleus is composed of not one, but three cyclohexane rings and one cyclopentane ring. Three fused cyclohexane rings also constitute adamantane, whose derivatives have been widely used in medical chemistry. Particularly, adamantane is found in antivirals and drugs against infectious diseases, in drugs to treat Central Nervous System illnesses or in cancer research.<sup>17</sup>



**Fig. 2** Examples of popular multiple cyclohexane motifs in natural products

According to a study carried out by Aldeghi and colleagues in 2014,<sup>18</sup> cyclohexane was number six in the list of ‘ten most frequent ring fragments found in marketed drugs’. Thus, cyclohexanes are found in drugs such as anaesthetics, like phencyclidine (later replaced with ketamine); antidepressants of the SNRI family such as venlafaxine and desvenlafaxine; or anticancer and chemotherapeutic drugs like docetaxel or etoposide.

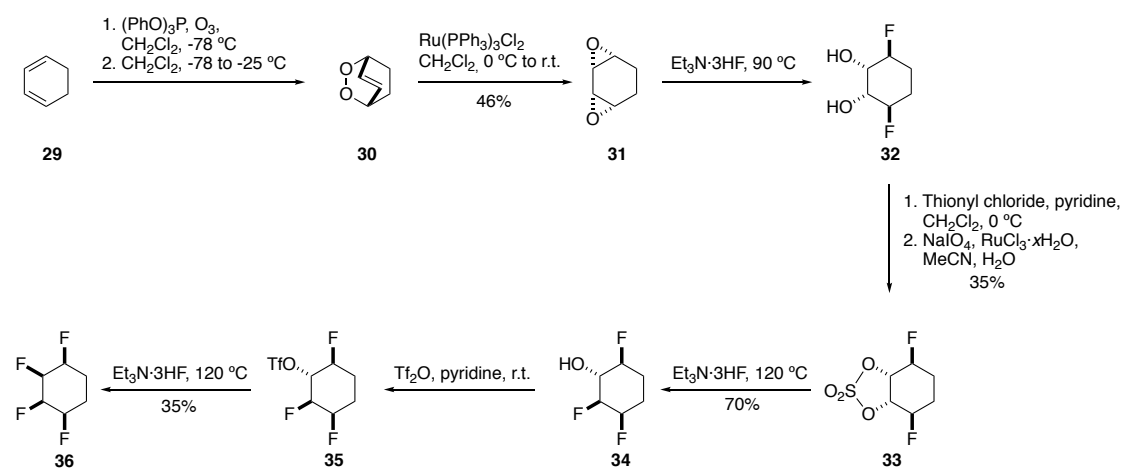


**Fig. 3** Examples of marketed drugs containing cyclohexane motifs

The substitution of hydrogen for fluorine atoms has had an important impact in drug development.<sup>1,2,3</sup> However, even though cyclohexane and fluorine are very commonly found in organic chemistry, examples in which they appear together, as fluorocyclohexanes, are very rare.

In an attempt to change the classical properties of cyclohexane, the St Andrews group started to explore all-*cis* vicinal fluorocyclohexanes, with the syntheses of all-*cis* 1,2,3,4-tetrafluorocyclohexane<sup>19</sup> (**Scheme 1**) and all-*cis* 1,2,4,5-tetrafluorocyclohexane<sup>20</sup> (**Scheme 2**) in 2011. The characteristic all-*cis* disposition of the fluorine atoms in these derivatives imparts exceptionally high dipole moments for aliphatic rings, making them interesting new motifs.<sup>19,20</sup>

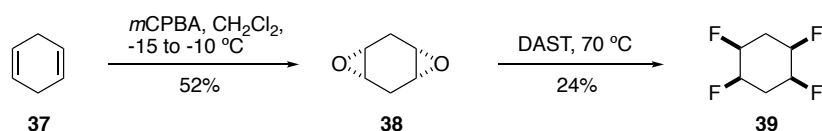
The synthesis of all-*cis* 1,2,3,4-tetrafluorocyclohexane **36** is shown in **Scheme 1**.<sup>19</sup> Stereochemical control was required for the formation of all-*cis* 1,2,3,4-tetrafluorocyclohexane **36**, sequentially converting the C-O bonds of *cis* 1,3-diepoxy cyclohexane **31** to C-F bonds.



**Scheme 1.** Previous synthesis for all-*cis* 1,2,3,4-tetrafluorocyclohexane<sup>19</sup>

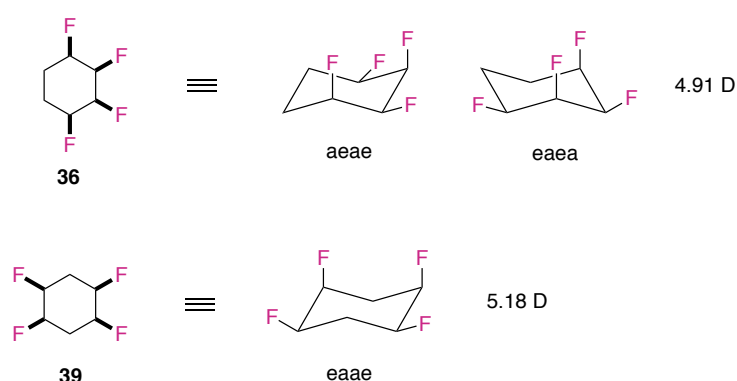
A preparation of the isomer all-*cis* 1,2,4,5-tetrafluorocyclohexane **39** was published shortly after (Durie *et al.*, 2012),<sup>20</sup> as shown in **Scheme 2**. The procedure was more direct and consisted of only two steps. Diepoxidation of 1,4-cyclohexadiene with *m*CPBA, followed by fluorination of the resultant diepoxide with DAST at 70 °C gave the product (**39**). This final step was low yielding, however overall the route proved practical to generate milligrams of this solid (m.p. = 107-109 °C).<sup>20</sup>





**Scheme 2.** Published synthesis for all-*cis* 1,2,4,5-tetrafluorocyclohexane<sup>20</sup>

A key characteristic of both of these structures is that they have 1,3-diaxial C—F bonds in the chair conformation, and ring interconversion also generates 1,3 diaxial conformers (**Fig. 4**). This special arrangement of the fluorines gives them high molecular dipole moments: 4.91 D for **36** and 5.18 D for **39**.<sup>19,20</sup>

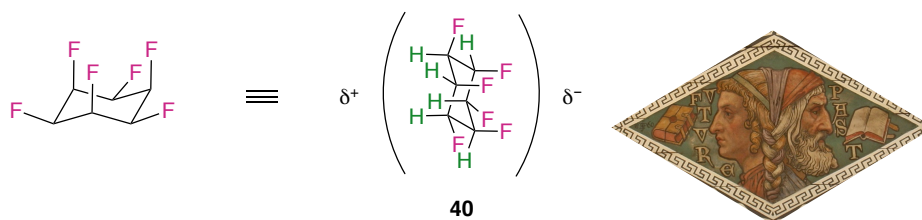


**Fig. 4** Chair conformations of **36** and **39**

This progressed to the synthesis of all-*cis* hexafluorocyclohexane **40** in St Andrews in 2015.<sup>21</sup> Cyclohexane **40** is the stereoisomer of 1,2,3,4,5,6-hexafluorocyclohexane with the highest energy, and one of the most polar aliphatic molecules known ( $\mu = 6.2$  D).<sup>21</sup>

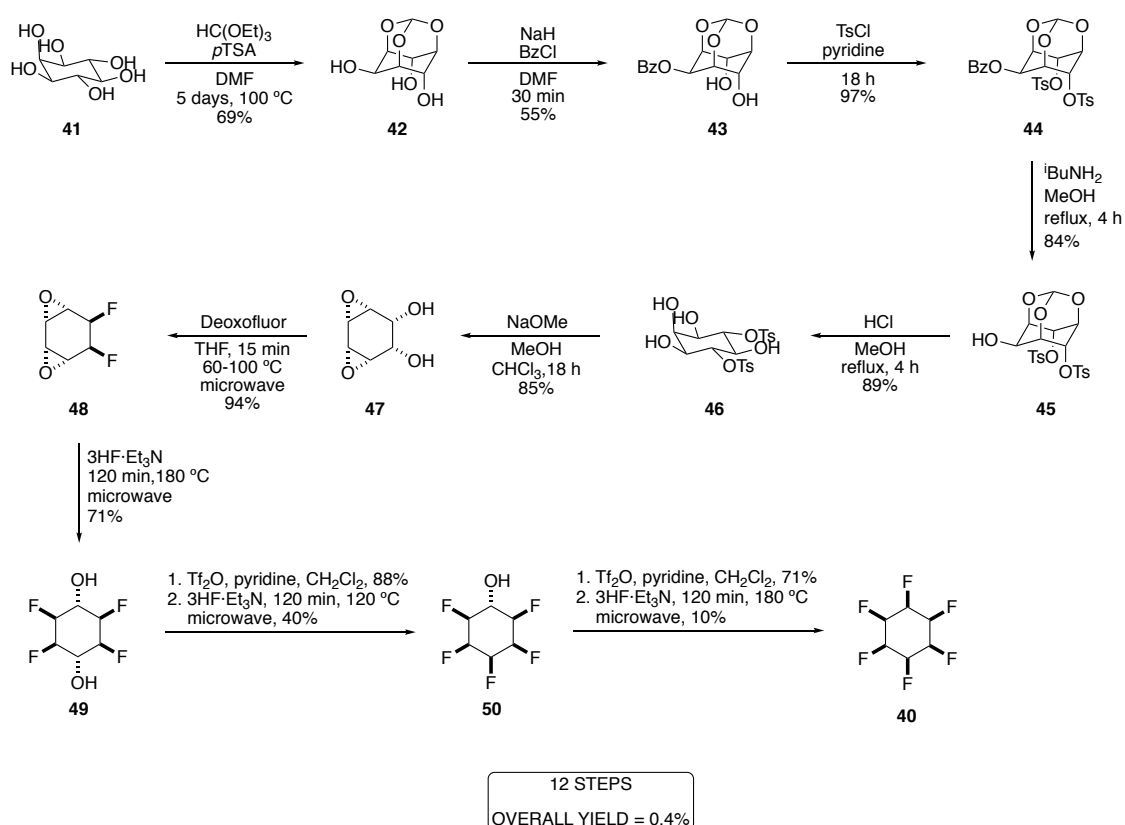
Following the St Andrews synthesis, a direct hydrogenation of hexafluorobenzene was developed by Glorius in Germany in 2017.<sup>22</sup>

In practice, this all-*cis* arrangement creates molecules with differentiated faces: an electronegative cation-binding fluorine face and an electropositive anion-binding hydrogen face.<sup>21</sup> They have been referred to as ‘Janus-faced’ compounds, given their resemblance to the Roman god of past and future Janus (usually represented with two faces: one overlooking the past, and one overlooking the future), as shown in **Fig. 5**.



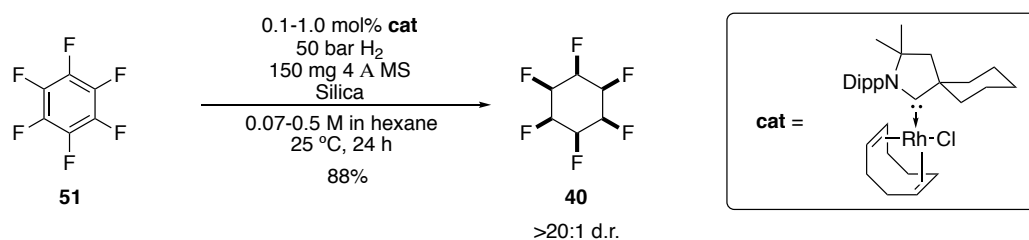
**Fig. 5** Janus-faced all-*cis* hexafluorocyclohexane

The St Andrews synthesis is shown in **Scheme 3**, and consists of 12 steps.<sup>21</sup>



**Scheme 3.** Synthesis for all-*cis* 1,2,3,4,5,6-hexafluorocyclohexane by Keddle *et al.*<sup>21</sup>

The last fluorination step is particularly limiting with a yield of approximately 10%. This synthetic route is not appropriate for large-scale production of **40**. The Münster synthesis is much more efficient, generating **40** in a 34% yield, in just one step.<sup>22</sup> A yield that was further improved to 88%, by the addition of silica gel.<sup>23</sup> This method uses a specific rhodium catalyst, as shown in **Scheme 4**.

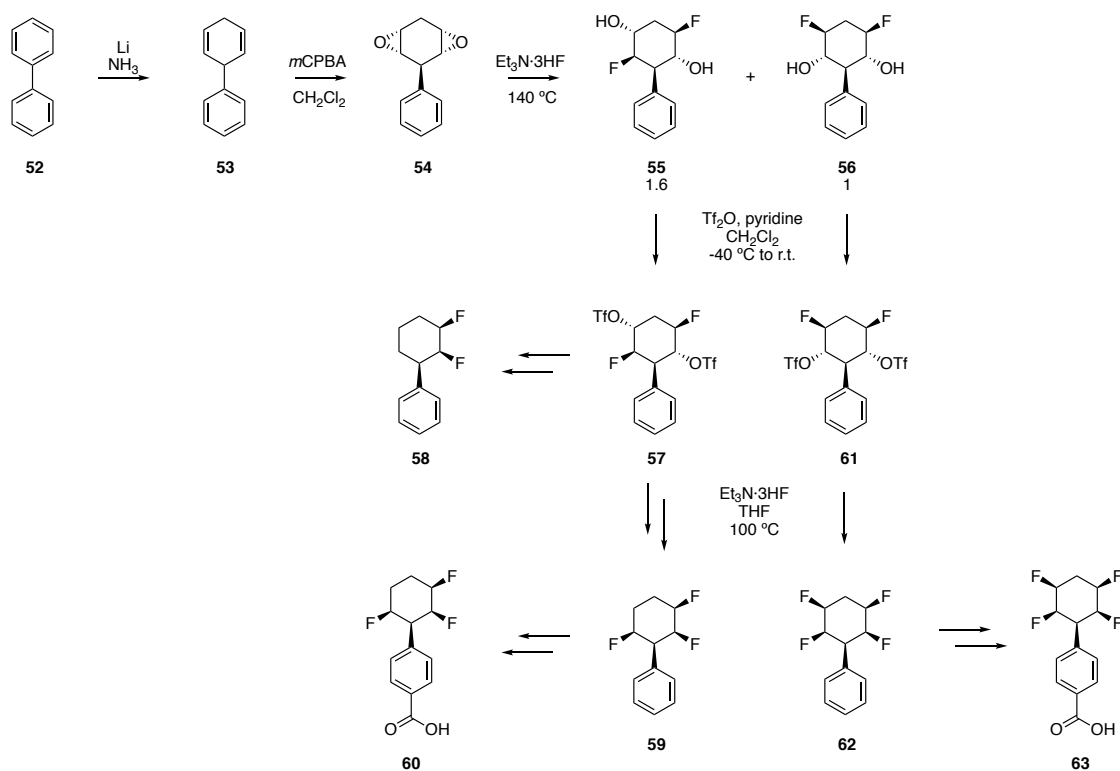


**Scheme 4.** One step synthesis for all-*cis* 1,2,3,4,5,6-hexafluorocyclohexane **40**<sup>23</sup>

Despite the significant improvement for the synthesis of fluorocyclohexanes, and particularly of all-*cis* 1,2,3,4,5,6-hexafluorocyclohexane **40** and both isomers of all-*cis* tetrafluorocyclohexane **36** and **39**, the generation of functionalised derivatives remains a challenge.

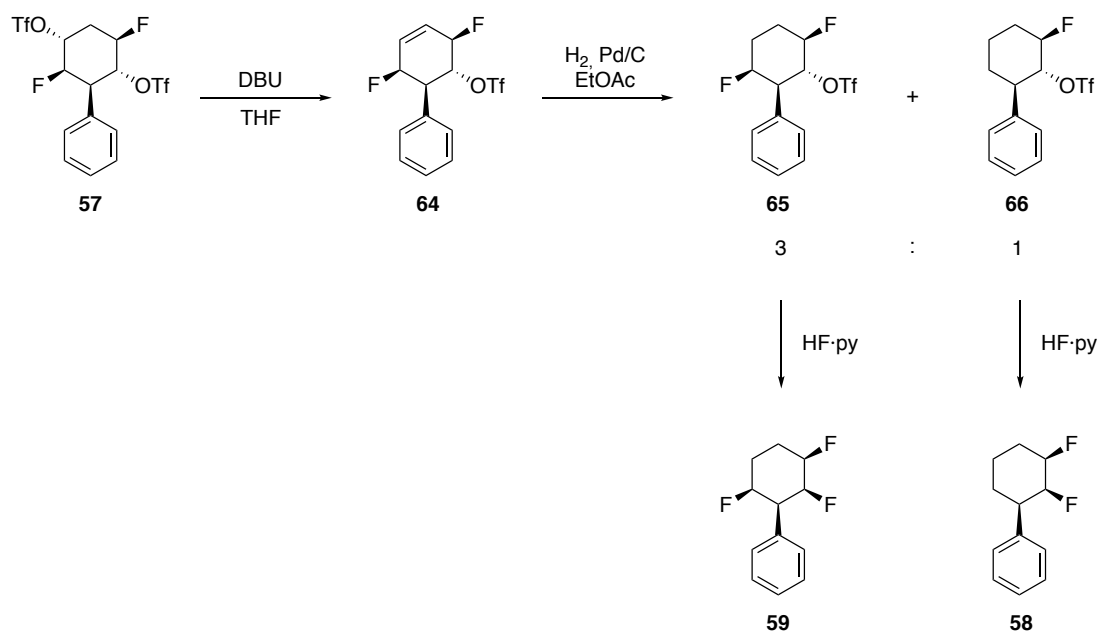
Currently the most developed approach in this regard are the phenyl derivatives of all-*cis* 1,2,4,5-tetrafluorocyclohexane **39**.<sup>24,25,26,27</sup> A secondary advantage deriving from the attachment of the phenyl ring is the higher volatility of the precursors to these products, which considerably facilitates their manipulation.

As shown in **Scheme 5**, the syntheses of the all-*cis* fluorocyclohexyl derivatives start from biphenyl, which is transformed to 1,4-dihydro-1,1'-biphenyl (**53**) by means of a Birch reduction. The resultant diene is then treated with *m*CPBA to give diepoxide **54** and then with hydrogen fluoride to generate the mixture of isomers **55** and **56** (1.6:1). Both isomers, once separated, are activated as their triflates and then fluorinated. These reactions lead to the tetra- and tri- fluorocyclohexyl derivatives (**62** and **59**), respectively.<sup>24,25,26,27</sup>



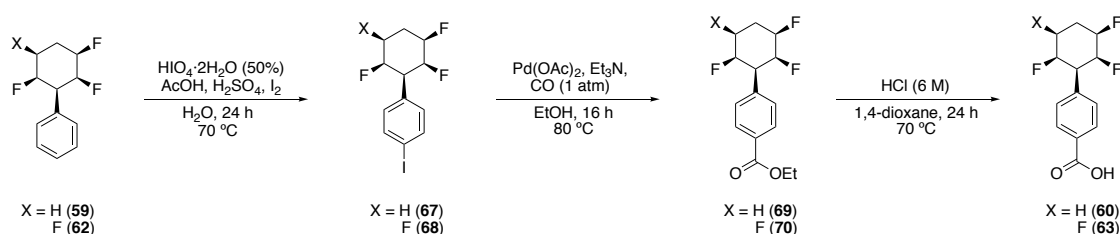
**Scheme 5.** Syntheses of derivatised fluorocyclohexanes <sup>24,25,26,27</sup>

**Scheme 6** shows the synthesis of the tri- and di- fluoro derivatives from the same intermediate (**57**). After triflate elimination with DBU, hydrogenation of the double bond gives cyclohexanes **65** and **66**, the latter with the loss of a fluorine atom. Once the hydrogenation products have been purified, they are both treated in the same manner with HF·pyridine, to give **59** and **58**, respectively. <sup>25,26</sup>



**Scheme 6.** Route to **59** and **58** from compound **57**<sup>25,26</sup>

Finally, the carboxylic acid derivatives **60** and **63** were accessed by a four-step synthesis involving iodination, palladium-catalysed carbonylation and hydrolysis of the esters **69** and **70** to give the corresponding carboxylic acids, **60** and **63** (Scheme 7).<sup>27</sup>

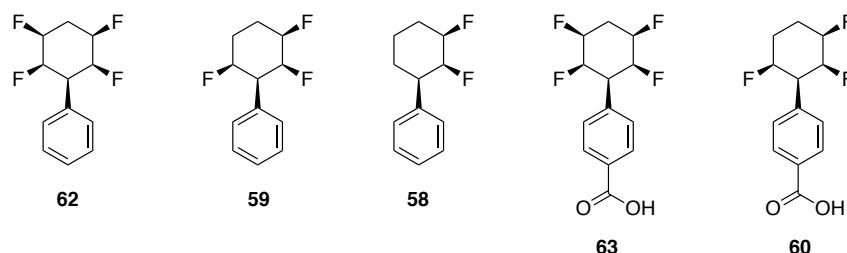


**Scheme 7.** Synthesis of carboxylic acids **60** and **63** from **59** and **62**<sup>27</sup>

Compounds **58**, **59**, **60**, **62** and **63** used in the following sub-sections of this Chapter were prepared by Dr Mohammed Salah Ayoup (University of St Andrews).

### 2.1.2 Aims and Objectives

In order to study how the number of fluorine atoms affects the metabolic stability of these derivatives, all-*cis* tetra-, tri-, and difluorocyclohexyl benzenes **62**, **59** and **58** were selected for this study, as well as the two *para*-substituted benzoic acid derivatives **60** and **63**.



**Fig. 6** Selected molecules for the metabolic studies with *C. elegans*

The fungus *Cunninghamella elegans* was used to explore the metabolic vulnerability of each of the selected compounds, and their resultant metabolites were isolated by HPLC prior to full characterisation by NMR spectroscopy and mass spectrometry.

Additionally, a reverse-phase HPLC method for the estimation of lipophilicities was designed and developed.

### 2.1.3 Estimation of lipophilicities by the use of $^{19}\text{F}$ NMR spectroscopy

Historically, the ‘drug-likeness’ of a molecule was assessed according to a series of factors, which are generally formalised as the ‘Lipinski’s Rule of Five’. According to Lipinski, a successful candidate would not have more than five hydrogen bond donors and ten hydrogen bond acceptors. The molecular weight would be less than 500 Da, and its octanol-water partition coefficient (expressed as Log *P*) should be lower than five.<sup>28,29</sup>

The traditional procedure to measure partition coefficients is the ‘shake-flask method’ between octanol and water. However, this method can be difficult to achieve reproducible results, and dissimilar values are often reported in literature.<sup>30</sup>

In 2016, at the start of this project, Linciau and coworkers published an accurate method of estimating the Log *P* values of particularly polar fluorinated compounds by

the use of  $^{19}\text{F}$  NMR spectroscopy.<sup>31</sup> This procedure is based on the ‘shake-flask method’, and consists of addition of the product, with an internal reference, to a biphasic medium of octanol and water. Then, a  $^{19}\text{F}$  NMR spectrum of each of the phases is recorded, and the peaks integrated relative to each other in the different phases.

The relationship between the integration values and the Log  $P$  estimations is as follows.<sup>31</sup>

First, it is necessary to define the integration ratio ( $\rho$ ) for the aqueous and the octanol phases. In the following equations, ‘X’ refers to the compound to be measured and ‘ref’ is the reference.

$$\rho_{\text{oct}} = \frac{I_{\text{oct}}^{\text{X}}}{I_{\text{oct}}^{\text{ref}}} = \frac{n^{\text{X}}C_{\text{oct}}^{\text{X}}}{n^{\text{ref}}C_{\text{oct}}^{\text{ref}}}$$

$$\rho_{\text{aq}} = \frac{I_{\text{aq}}^{\text{X}}}{I_{\text{aq}}^{\text{ref}}} = \frac{n^{\text{X}}C_{\text{aq}}^{\text{X}}}{n^{\text{ref}}C_{\text{aq}}^{\text{ref}}}$$

Once the integration ratio for both phases has been defined, it is possible to calculate the partition coefficient ( $P$ ), can be calculated as follows:

$$\frac{\rho_{\text{oct}}}{\rho_{\text{aq}}} = \frac{C_{\text{oct}}^{\text{X}}C_{\text{aq}}^{\text{ref}}}{C_{\text{aq}}^{\text{X}}C_{\text{oct}}^{\text{ref}}} = \frac{p^{\text{X}}}{p^{\text{ref}}}$$

$$p^{\text{X}} = p^{\text{ref}} \left( \frac{\rho_{\text{oct}}}{\rho_{\text{aq}}} \right)$$

$$\log p^{\text{X}} = \log p^{\text{ref}} + \log \left( \frac{\rho_{\text{oct}}}{\rho_{\text{aq}}} \right)$$

Since the log  $P$  of the reference is well known, the only unknown variable is the log  $P$  of the compound under investigation.

For this NMR method, the settings of both the actual NMR experiment and that for processing the FID are extremely relevant, such that the results are reproducible. According to the method, the important parameters are the frequency offset point (OIP) and the pulse delay time (D1). In order to obtain the best signal to noise ratio – and therefore a more accurate integration – the pulse delay time (D1) of the samples needs to be sufficiently long. These were evaluated by Linclau such that D1 for

octanol should be 30 s and for water, 60 s. The frequency offset point (O1P) should be centred on the fluorine signals, thus it varies for each sample. Other variables were studied but found to be less relevant, such as spectral width (SD) or the number of scans (NS).<sup>31</sup> Therefore, unless otherwise stated, the number of scans was left on the automatic selection for the instrument (NS = 64 scans). All spectra were processed in the same way, using Topspin 3.5p15.

#### **2.1.3.1 Calibration Experiments**

The necessity for some calibration experiments became apparent at the start of this project, in order to adapt the method to the selected compounds. To probe the suitability and reproducibility of the procedure, it was decided to run two sample experiments, with two molecules whose Log *P* values had been previously evaluated against 2,2,2-trifluoroethanol (TFE, **71**): 1,3-difluoro-2-propanol (**72**) and 2-fluorophenol (**73**).<sup>31</sup> In his article, Linclau used both as references, depending on the chemical shift of the products to be assessed.<sup>31</sup>

##### **2.1.3.1.1 2,2,2-Trifluoroethanol (71) against 1,3-difluoro-2-propanol (72)**

2,2,2-Trifluoroethanol (TFE, **71**) was chosen as the first reference compound, to estimate the value of 1,3-difluoro-2-propanol's Log *P*. This exact experiment was published by Linclau, and therefore allows the reproducibility to be assessed.

The procedure and sample preparation followed the published method,<sup>31</sup> by adding TFE and 1,3-difluoro-2-propanol into a pear-shaped flask containing HPLC grade octanol and a stirrer bar. Finally, milliQ water was added. The flask was stirred for 3 hours at room temperature, and the phases were left to separate overnight.

The NMR samples were prepared carefully to avoid contamination between the phases. A fraction of the upper octanol layer was extracted with a 1 mL syringe and a needle, and some was discarded to ensure that no aqueous residues were left in the needle. The remaining sample were transferred into an NMR tube with d<sub>6</sub>-acetone, and homogenised.



The preparation of the lower aqueous sample required precaution to avoid octanol contamination. In this case, some air was taken with the syringe and needle and it was slowly discharged while passing through the organic layer to avoid accidentally collecting any octanol. The remaining air was slowly discharged into the aqueous phase to avoid any mixing, and a sample of the aqueous layer was taken. The needle and syringe were removed as efficiently as possible, and some water was discarded to avoid contamination. The needle was also externally wiped with a tissue. The remaining sample was added into an NMR tube followed by d<sub>6</sub>-acetone, and homogenised.

In both cases, a full range <sup>19</sup>F {<sup>1</sup>H} NMR spectrum was run. Manual FID processing led to the following results:

$$\rho_{\text{water}} = \frac{I^{72}}{I^{71}} = \frac{100.00}{60.24} = 1.62$$

$$\rho_{\text{octanol}} = \frac{I^{72}}{I^{71}} = \frac{27.08}{100.00} = 0.271$$

$$\log P^{71} = +0.36$$

$$\log P^{72} = \log P^{71} + \log\left(\frac{\rho_{\text{oct}}}{\rho_{\text{water}}}\right) = +0.36 + \log\left(\frac{0.271}{1.62}\right) = +0.36 + (-0.79) = -0.43$$

This value matched the one described by Linclau ( $\log P^{72} = -0.42$ ), however, the method of integration is important, since the automated integration readout was inaccurate:

$$\rho_{\text{water}} = \frac{I^{72}}{I^{71}} = \frac{0.9993}{0.3405} = 2.94$$

$$\rho_{\text{octanol}} = \frac{I^{72}}{I^{71}} = \frac{0.3362}{1.0000} = 0.3362$$

$$\log P^{71} = +0.36$$

$$\log P^{72} = \log P^{71} + \log\left(\frac{\rho_{\text{oct}}}{\rho_{\text{water}}}\right) = +0.36 + \log\left(\frac{0.336}{2.94}\right) = +0.36 + (-0.94) = -0.58$$

In **Table 1**, the different values obtained for 1,3-difluoro-2-propanol are displayed, in order to establish a comparison.

**Table 1.** Comparison of Log *P* values for 1,3-difluoro-2-propanol from different sources

Source	Log <i>P</i> value
Literature (calculated, cLog <i>P</i> )*	-0.36
Linclau (experimental) <sup>31</sup>	-0.42
This project (experimental, automatic integration)	-0.58
This project (experimental, manual integration)	-0.43

\* The cLog *P* of -0.36 was calculated by an ACD/Labs Software V11.02 (© 1994-2018 ACD/Labs), and it does not match experimental values. The literature value was also retrieved from the U.S. National Library of Medicine online database (<http://chem.sis.nlm.nih.gov/chemidplus/>)

With these results, it became evident that the method for integrating the signals is extremely important, given that a very small change provokes a significant difference in the final value – even within the same experiment. It is also necessary to repeat experiments to ensure reproducibility.

#### 2.1.3.1.2 TFE (71) against 2-fluorophenol (73)

The second calibration experiment was carried out with two objectives. The first was to ensure the reproducibility of the method; and the second, to obtain an experimental value for 2-fluorophenol. 2-Fluorophenol (**73**) has a more similar chemical shift to the all *cis*-fluorocyclohexyl motifs than TFE, and therefore, its use as a reference compound should lead to more accurate Log *P* evaluations.

This experiment was conducted by reproducing the conditions used in the previous calibration test. TFE and 2-fluorophenol were added into a pear-shaped flask with octanol and MilliQ distilled water, and stirred for 3 hours.

Again, the way the FID processing was carried out was crucial, leading to different log *P* values when automatic and manual integration methods were applied. **Table 2** shows the different values obtained for 2-fluorophenol depending on the conditions used, and those described in the literature.

**Table 2.** Comparison of Log *P* values for 2-fluorophenol from different sources

Source	Log <i>P</i> value
Literature (cLog <i>P</i> ) <sup>30b</sup>	1.82
Linclau (experimental) <sup>31</sup>	1.70
This project (experimental, automatic integration)	1.72
This project (experimental, manual integration)	1.70

For a second time, the manual integration showed better consistency with Linclau's results. This time, the  $\Delta\text{Log } P = 0.02$  between the automated and the manual results, was much closer. After these calibration experiments, 2-fluorophenol **73** was considered as a good reference compound.

#### **2.1.3.2 Log *P* estimation of the selected all-*cis* fluorocyclohexanes**

##### **2.1.3.2.1 All-*cis* tetrafluorocyclohexylbenzene (**62**)**

The chosen reference 2-fluorophenol (**73**) was used to determine a lipophilicity value for all-*cis* tetrafluorocyclohexylbenzene **62**, using an analogous procedure to that of the calibration assays.

NMR samples were prepared as usual, and run using the following settings: aqueous sample (NS = 64 scans, D1 = 60 s, O1P = -169 ppm); octanol sample (NS = 64 scans, D1 = 30 s, O1P = -169 ppm). The centre of the spectrum, established by the frequency offset point (O1P), was set between the reference signal and that of compound **62**, to ensure the best accuracy. Manual integration of the peaks led to a Log *P* value of 2.45.

The experiment was repeated, and a Log *P* value of 2.46 was obtained, thus indicating good reproducibility ( $\Delta\text{Log } P = 0.01$ ) for both runs. After these results, the method was understood to be reliable if manual integration is used. However, the intensity of the NMR signals was quite low, given the very poor solubility of **62** in water or octanol.

#### 2.1.3.2.2 All-*cis* trifluorocyclohexylbenzene (**59**)

The treatment of all-*cis* trifluorocyclohexylbenzene (**59**) with 2-fluorophenol in octanol and water as described allowed NMR analysis (O1P = -166 ppm) of the phases, leading to a Log *P* value of 2.36. Repetition of the assay afforded a Log *P* value of 2.38.

The close difference between these two experiments ( $\Delta\text{Log } P = 0.02$ ), suggests good reproducibility. However, the increasing error when compared with the tetrafluoro derivative **62** can be explained by the lower solubility of **59** in water, producing a lower signal to noise ratio, and influencing the integration values.

#### 2.1.3.2.3 All-*cis* difluorocyclohexylbenzene (**58**)

The procedure followed for this experiment was identical to those previously described. However, in this case, the very low intensity of the fluorine signals (O1P = -168 ppm) in the aqueous phase did not allow the determination of a Log *P* value for **58**. The NMR experiment for the aqueous sample was increased to 128 scans instead of 64, however this did not improve the signal to noise ratio.

A second assay was carried out, in order to establish whether it was possible to obtain a better NMR signal of the aqueous phase containing **58**. To this end, the amount of **58** was increased to 10 mg, and the amount of octanol decreased to 1.5 mL, to favour its solubility in water. Stirring was extended for 6 hours before leaving the phases to separate. However, this did not improve the concentration of **58** in the aqueous phase, as determined by NMR spectroscopy.

#### 2.1.3.2.4 All-*cis* tetrafluorocyclohexyl carboxylic acid (**63**)

Cyclohexane **63** was available in very low amounts, given its challenging synthesis in the lab. Therefore, only one experiment to determine its lipophilicity was carried out. Examination of NMR spectra (O1P = -170 ppm) allowed a preliminary estimation of the Log *P* value as 1.36. Additionally, the solubility of **63** in water is extremely low, and determination of the integral values for the  $^{19}\text{F}$  NMR signals was unreliable.

#### 2.1.3.2.5: All-*cis* trifluorocyclohexyl carboxylic acid (**60**)

The assay to determine the lipophilicity value of **60** presented the same challenges as that of the previous experiment. While scarcity of product allowed only one experiment, its low solubility did not allow a precise manual integration. Hence, the recorded value ( $\text{Log } P = 1.46$ ) is only a preliminary one.

### Conclusions on the use of $^{19}\text{F}$ $\{^1\text{H}\}$ NMR for the estimation of $\text{Log } P$ values

When possible, the experiments in this section were repeated at least once to ensure reproducibility.

Despite the convenience of the  $^{19}\text{F}$  NMR method, it was not particularly successful for the (fluorocyclohexyl)benzene derivatives, which have  $\text{Log } P$  values higher than 2. The poor solubility of this family of compounds – especially in the aqueous phase – gave poor signal to noise ratios, which compromised the accuracy of the  $\text{Log } P$  values obtained.

Thus, although the method described by Linclau<sup>31</sup> is good for fluorinated sugars and other compounds of high water solubility (and low  $\text{Log } P$ ), it is limited in terms of compounds with  $\text{Log } P$ s in the 2 to 5 range.

### 2.1.4 Determination of $\text{Log } P$ values by reverse-phase HPLC

An alternative method for the estimation of the  $\text{Log } P$  values was explored, this time by reverse-phase HPLC, based on retention time on a  $\text{C}_{18}$  column with an aqueous phase. The estimation of lipophilicity values was conducted using a Phenomenex Luna  $\text{C}_{18}$  100A (250 mm x 4.60 mm)  $5\mu$  column in a Shimadzu Prominence HPLC, based on the premise that  $\text{Log } P$  values are directly proportional to retention times, as stated by Gianinis in 2008.<sup>32</sup>

A series of reference compounds, with known literature lipophilicity values,<sup>30</sup> were evaluated (injecting 5-10  $\mu\text{L}$  of 0.5 mg/mL solution in MeCN). The retention times ( $R_t$ ) were used to calculate the capacity factors ( $k$ ) with the following equation:

$$\text{Capacity factor (k)} = \frac{\text{Retention Time} - \text{Dead Time of the column}}{\text{Dead Time of the column}}$$

The dead time of the column is the time that takes for an unretained molecule (such as the solvent) to pass through. For the particular experimental conditions used, the dead time of the column was calculated to be 1.97 min.

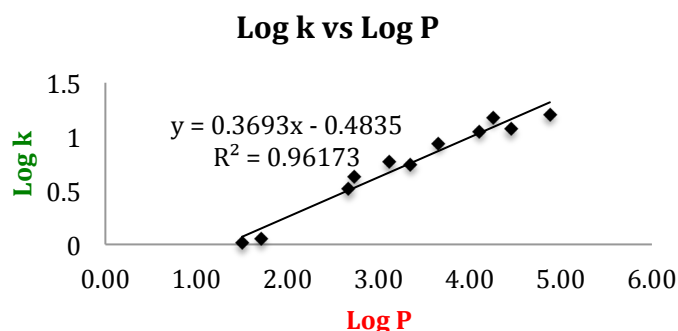
**Table 3** shows the experimental values obtained and the calculations used to evaluate Log *P*s. Literature Log *P* values<sup>30</sup> of the reference compounds are displayed in red, along with the experimental retention times observed for each molecule (measured in triplicate). Average retention times were calculated using each of the three values obtained, and this average value was used to calculate the capacity factor for each reference compound (using the equation described above). Finally, the logarithm of *k* was calculated (values shown in green).

For this study, the experimental conditions (60:40 MeCN:water, 1mL/min) included supplementation of the eluents with 0.05% of TFA, in order to protonate the carboxylic acids, and obtain sharper peaks. The set of conditions shown in **Table 3** are the result of multiple optimisations of the method, developed in order to give the most reproducible values.

**Table 3.** Literature Log *P*,<sup>30</sup> measured retention time (*R<sub>t</sub>*) and capacity factors (*k*) for the references

Reference	Log <i>P</i>	Rt 1 (min)	Rt 2 (min)	Rt 3 (min)	Average Rt (min)	Capacity factor ( <i>k</i> )	Log <i>k</i>
Phenol	1.50	3.99	3.99	3.99	3.99 ± 0.00	1.03	0.013
2-Fluorophenol	1.71	4.18	4.18	4.18	4.18 ± 0.00	1.13	0.051
Benzofuran	2.67	8.42	8.42	8.43	8.42 ± 0.01	3.29	0.52
Toluene	2.73	10.27	10.27	10.27	10.27 ± 0.00	4.23	0.63
<i>o</i> -Xylene	3.12	13.36	13.34	13.37	13.36 ± 0.02	5.80	0.76
Naphthalene	3.35	12.75	12.75	12.78	12.76 ± 0.02	5.50	0.74
Cumene	3.66	18.87	18.88	18.91	18.89 ± 0.02	8.61	0.94
<i>t</i> -Butylbenzene	4.11	23.93	23.86	23.91	23.90 ± 0.04	11.16	1.05
Butylbenzene	4.26	31.10	31.15	31.14	31.13 ± 0.03	14.84	1.17
Anthracene	4.45	25.08	25.07	25.02	25.06 ± 0.04	11.75	1.07
Pyrene	4.88	33.24	33.33	33.17	33.25 ± 0.08	15.92	1.20

The calculated log *k* (Y-axis) values were plotted against their reported Log *P* values (X-axis) to obtain a linear regression equation, as represented in **Plot 1**.



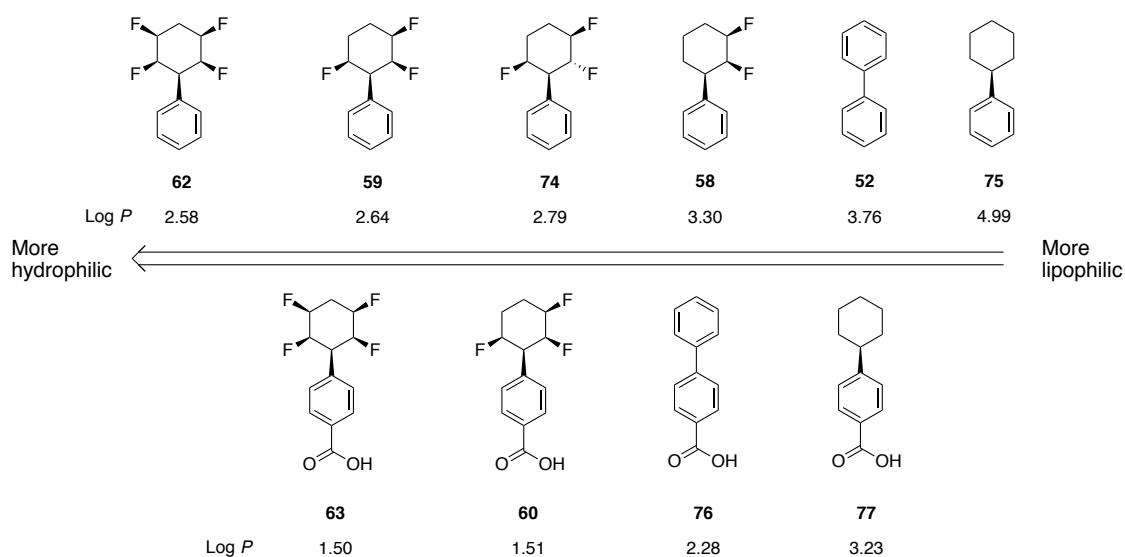
**Plot 1.** Regression line obtained for the reference compounds

The equation obtained in **Plot 1** was used to calculate the Log *P* values of a series of compounds by substitution of the logarithm of their capacity factor values in the equation above. Retention times, capacity factors (*k*) and their logarithms are shown in **Table 4**, along with the Log *P* values (blue), estimated by substitution of the Log *k* values in the regression equation in **Plot 1**.

**Table 4.** Estimation of Log *P* (blue) by using the logarithm of the measured capacity factors (green)

Compound	Log <i>P</i>	Rt 1	Rt 2	Rt 3	Average Rt	<i>k</i>	Log <i>k</i>
62	2.58	7.78	7.78	7.72	7.76 ± 0.04	2.95	0.47
59	2.64	8.08	8.08	8.03	8.06 ± 0.03	3.10	0.49
63	1.50	4.28	4.27	4.28	4.28 ± 0.01	1.18	0.07
60	1.51	4.29	4.30	4.30	4.30 ± 0.01	1.19	0.07
58	3.30	12.68	12.68	12.61	12.66 ± 0.05	5.44	0.74
74	2.79	8.87	8.88	8.87	8.87 ± 0.01	3.52	0.55
52	3.76	17.82	17.81	17.71	17.78 ± 0.07	8.05	0.91
75	4.99	46.88	46.53	46.89	46.77 ± 0.24	22.80	1.36
76	2.28	6.45	6.45	6.45	6.45 ± 0.00	2.28	0.36
77	3.23	12.00	11.99	12.00	12.00 ± 0.01	5.11	0.71

**Table 4** also presents cyclohexyl and phenyl analogues of the studied compounds which were selected in order to establish lipophilicity comparisons, particularly to explore the effect of the 1,3-diaxial C—F bonds on the polarity. The effect of the number of fluorine atoms and the presence of the carboxylic acid group were also probed. For a more visual representation, structures are shown by families and in increasing lipophilicity order in **Fig. 7**.



**Fig. 7** Comparison of the Log *P* values and lipophilicity trends for selected fluorocyclohexanes and comparative analogues

The lowest Log *P* (most hydrophilic) values are those for the tetrafluorinated analogues, followed by the trifluorinated and difluorinated compounds, while the structures without fluorine display the highest lipophilicities. It is clear that the number of fluorine atoms affects the lipophilicity values considerably.

A comparison of **74** (which lacks 1,3-diaxial fluorines) with **59**, shows that the diaxial fluorine atoms do have an effect on lipophilicity, even though not as significant as the complete removal of a fluorine, as is the case of **58**. Hence, the increasing lipophilicity order goes from all-*cis* trifluorinated cyclohexane **59**, followed by the non-*cis* **74** to the difluorinated structure **58**, having **58** the highest lipophilicity of the fluoro series.

Carboxylic acids **60** and **63** are inherently more polar than the phenyl derivatives, however the increasing number of fluorines leads to a progressive lowering of their Log *Ps*.

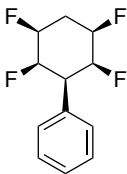
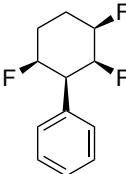
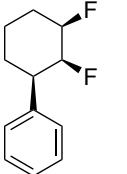


#### 2.4.1.1 Conclusions on the use of a reverse-phase HPLC method to estimate lipophilicities

The reverse-phase HPLC method for the estimation of lipophilicity values was very convenient and easy to use, although it required optimisation before obtaining a reliable set of values. Aqueous solubility does not constitute a problem, since the samples are prepared in acetonitrile, instead of an octanol-water mixture.

The values presented were not measured at neutral pH. TFA was added to acidify the eluent solvents, in order to measure the carboxylic acids. However, the values of the phenyl derivatives were also recorded at neutral pH (with no TFA present) to study their variation with pH. They did not vary significantly for the non-carboxylated products (**Table 5**), indicating the method is quite robust.

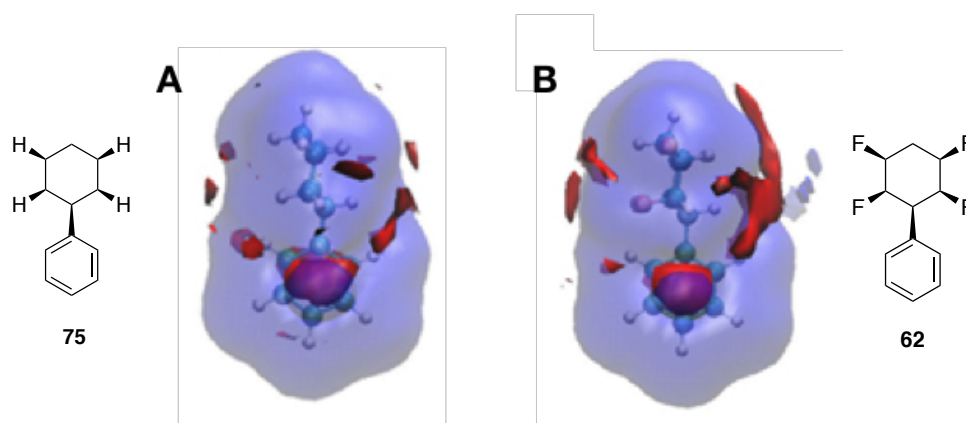
**Table 5.** Comparison of Log *P* values with and without supplemented TFA

	 <b>62</b>	 <b>59</b>	 <b>58</b>
<b>Log <i>P</i> (No TFA)</b>	2.54	2.60	3.24
<b>Log <i>P</i> (With TFA)</b>	2.58	2.64	3.30

A clear outcome from this work has been that not only does fluorine addition lower the lipophilicity of cyclohexane rings, but that the lipophilicity values are below Log *P* = 3.5, bringing these compounds into the range as candidates for drug scaffolds.

### 2.4.1.2 Molecular dynamics

Molecular dynamic studies evaluating how water interacts with the cyclohexyl derivatives were performed in collaboration with Stefano Bosisio and Julien Michel (University of Edinburgh).<sup>26,33</sup> The preferred location of water and the contribution of water molecules to enthalpy, entropy and free energy values was studied.<sup>26</sup> In this sense, **Fig. 8** illustrates a hydration profile for all-*cis* tetrafluoro cyclohexane **62** and its non-fluorinated analogue **75**.<sup>26</sup>



**Fig. 8** Average interactions of water molecules (red) with non-fluorinated **75** (structure A) and tetrafluorinated **62** (structure B)<sup>26</sup>

For both **75** and **62**, there is a similar association of water to each face of the phenyl ring, however, the cyclohexyl rings are different. Water density around the non-fluorinated cyclohexyl motif in **75** is quite low (A), while for in the fluorinated analogue **62**, it is significant (B).<sup>26</sup>

Given the opposite polarities of the faces in **62**, water associates most significantly with the hydrogen face and less significantly with the fluorine face. The polarity of the ring arises from the polarised hydrogens, and therefore, hydrogen bonding to water rationalises the increased hydrophilicity of the hydrogen face.<sup>26</sup>

### 2.1.5 Metabolic fate in *Cunninghamella elegans*

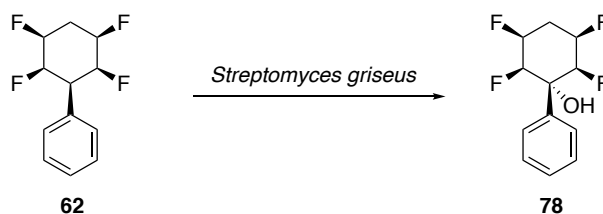
Given that the lipophilicity values for these cyclohexanes is within the medicinal chemistry range, a study of the metabolism of all-*cis* (fluorocyclohexyl)benzenes became an objective. This was carried out partly in collaboration with Prof. Cormac Murphy, at University College Dublin, including a three-month placement. The model organism *Cunninghamella elegans* was used.<sup>34,35</sup>

The incubation, inoculation and extraction procedures of compounds **62**, **59**, **58**, **63** and **60** in the fungus were carried out in a standardised manner. Compounds (as a solution in DMF) were added to mature liquid cultures of the fungus, incubated for 3 days, and extracted into ethyl acetate. Purification was achieved using reverse-phase HPLC, and the compounds characterised by NMR spectroscopy and mass spectrometry. A full description of the procedure can be found in the Experimental Chapter of this thesis.

Additionally, and in order to ensure that the biotransformations were carried out by the fungus, and not as result of decomposition from being in the medium for three days, blanks with no *C. elegans* were run for all the compounds. In these experiments, only starting materials were recovered in all cases.

#### 2.1.5.1 All-*cis* (2,3,5,6-tetrafluorocyclohexyl)benzene (**62**)

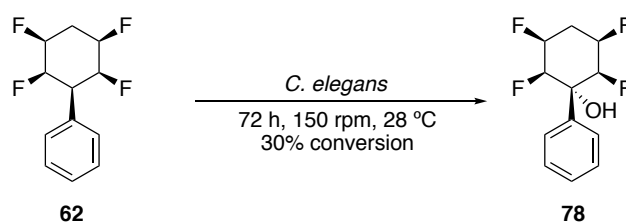
An earlier one off study on the metabolic fate of all-*cis* (2,3,5,6-tetrafluorocyclohexyl)benzene **62** had been investigated in *Streptomyces griseus* by Dr Mohammed S. Ayoup (University of St Andrews). Dr. Ayoup observed the formation of **78** as the only metabolite.



**Scheme 8.** Bioconversion of **62** to **78** (Dr. Ayoup, University of St Andrews)

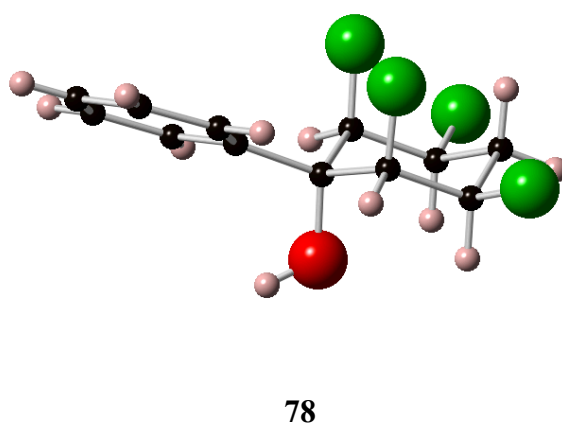
In that experiment, it was obvious after observation of the bacterial culture that there was some form of fungal contamination present. When the experiment was repeated using an aseptic batch of *S. griseus*, only starting material was recovered. These results seemed to indicate that the fungal contamination in the original culture was the cause of the biotransformation.

In this study, compound **62** was incubated with *C. elegans*. The experiment was carried out in triplicate to establish the reproducibility of the assays. Preliminary GC-MS analysis of the extracts indicated the formation of a new product, compatible with monohydroxylation at around 30% conversion.



**Scheme 9.** Metabolism of all-*cis* (tetrafluorocyclohexyl)benzene **62** in *C. elegans*

Reverse-phase HPLC purification allowed the isolation of starting material ( $t_{\text{ret}} = 23$  min) and this new metabolite ( $t_{\text{ret}} = 18$  min).  $^1\text{H}$  NMR spectroscopy and high-resolution mass spectrometry were consistent with monohydroxylation, and a crystal structure was obtained, confirming alcohol **78** as the product.

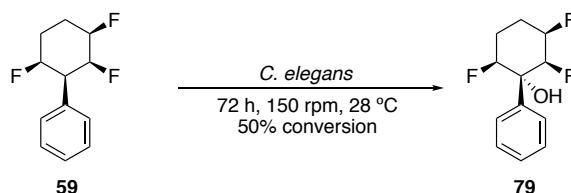


**Fig. 9** Crystal structure obtained by Dr. Ayoup for **78** <sup>26</sup>

#### 2.1.5.2 All-*cis* (2,3,6-trifluorocyclohexyl)benzene (**59**)

The influence of the number of fluorine atoms on the level of metabolism of the cyclohexanes was of interest. Therefore, it became necessary to gradually decrease the fluorine atoms present. Accordingly, all-*cis* (2,3,6-trifluorocyclohexyl) benzene **59** was probed.

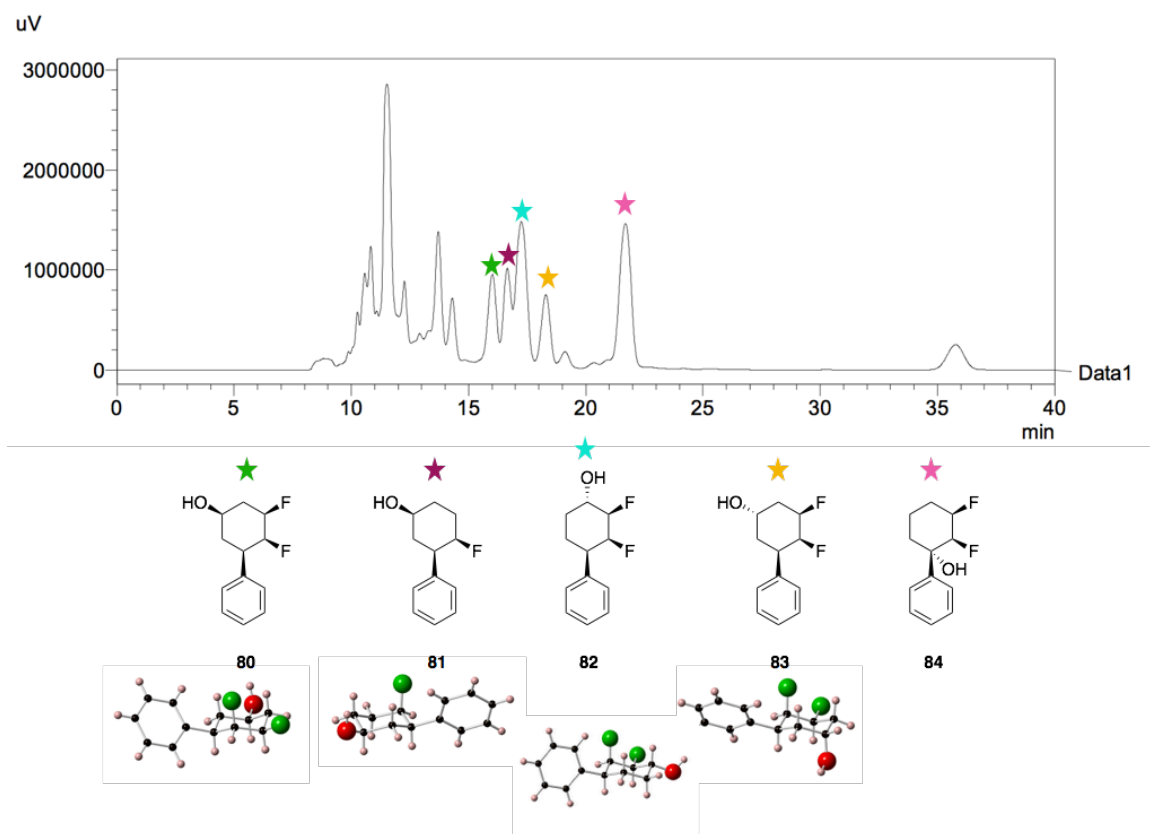
The experiments were carried out in triplicate and led to the same outcome, with the formation of one new metabolite, at a conversion of around 50% after 3 days. Full NMR and mass spectrometry analysis of the isolated metabolite indicated the benzyl hydroxylated product **79**.



**Scheme 10.** Metabolism of all-*cis* (2,3,6-trifluorocyclohexyl) benzene **59** in *C. elegans*

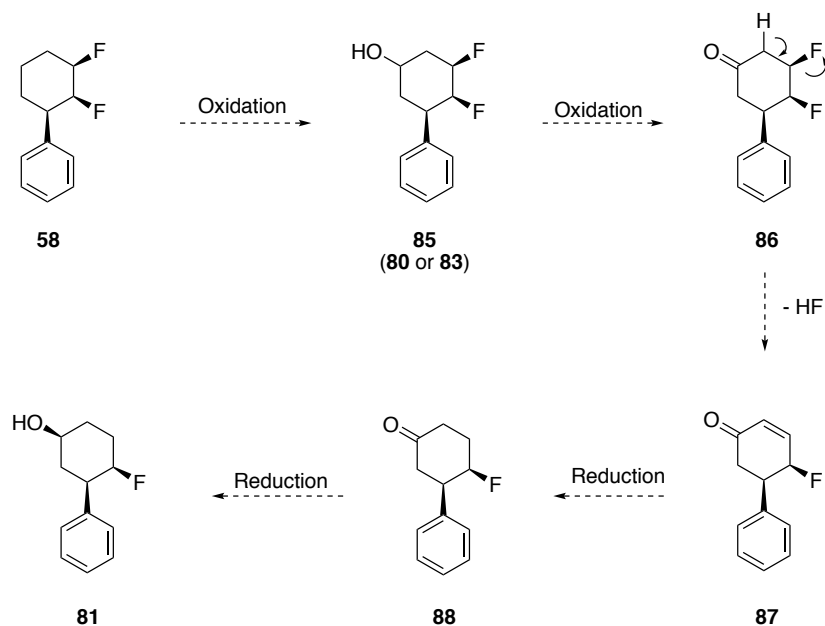
#### 2.1.5.3 All-*cis* (2,3-difluorocyclohexyl)benzene (**58**)

Incubation of **58**, which has two fluorine atoms, with *C. elegans* gave a complex mixture of products. Purification of these metabolites by reverse-phase HPLC proved to be challenging, given their close retention times. However, five different fluorinated metabolites were isolated. No starting material was recovered from these cultures, indicating complete conversion; and thus, suggesting the metabolic instability of this difluorinated derivative.



**Fig. 10** HPLC trace (70% AcCN:Water) for the extract, after incubating **58** with *C. elegans*

As shown in **Fig. 10**, the major metabolite remains the monohydroxylated product **84** (pink star), but other products are the result of *anti* hydroxylation in C4 (blue star, **82**) and of both *syn* and *anti* monohydroxylation at C5 (green and yellow stars, products **80** and **83**). The least anticipated product comes at  $t_{\text{ret}} = 16.5$  min (purple star, **81**). This metabolite features *syn* monohydroxylation at C5, relative to the fluorine position, and indicates the loss of a fluorine atom. The formation of this product would appear to require the action of more than one enzyme, and a pathway is suggested in **Scheme 11**.



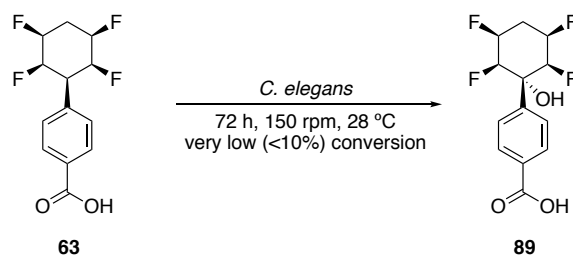
**Scheme 11.** Putative pathway for the formation of monofluorinated metabolite **81** in *C. elegans*

This hypothetical route starts with the observed metabolites **80** or **83** (stereoisomers), represented as single non-stereospecific structure **85** in **Scheme 11**. From here, oxidation to a ketone, followed by elimination of HF would offer a logical pathway. The conjugated double bond could then be reduced, and finally, the ketone converted to an alcohol.

In order to probe this pathway, both diastereoisomers **80** and **83** were inoculated into fungal cultures and incubated for 72 h, but none of these assays led to the formation of **81**. Metabolites **82** and **84** were also incubated with *C. elegans*, but again no generation of **81** was observed. Presumably, these metabolites are exuded from the cells and do not re-enter. Thus, the origin of the monofluorinated cyclohexane remains ambiguous.

#### 2.1.5.4 All-*cis* 4-(2,3,5,6-tetrafluorocyclohexyl)benzoic acid (**63**)

To further probe tetrafluorocyclohexane metabolism, incubation of all-*cis* 4-(2,3,5,6-tetrafluorocyclohexyl)benzoic acid **63** was explored. This gave one product, which once again was the result of benzylic hydroxylation. The yield was low and the product was obtained in less than 10% conversion from the starting material.

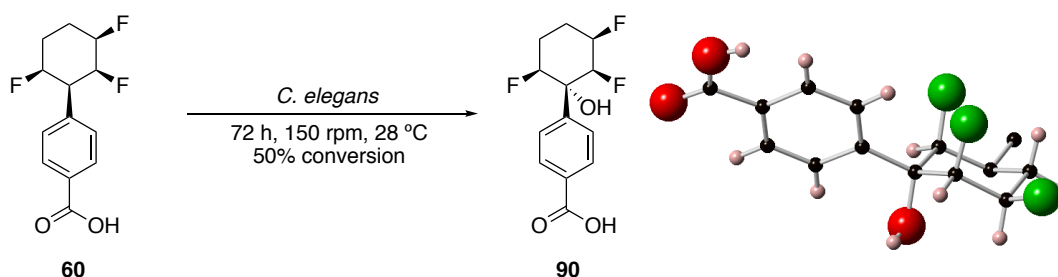


**Scheme 12.** *C. elegans* incubation of **63** gave **89**

The low amount of product recovered from the culture did not allow for full NMR characterisation. However, it is possible to follow the disappearance of the benzylic hydrogen atom by  $^1\text{H}$  NMR spectroscopy. A high-resolution accurate mass was obtained for product **89**, and was consistent with the structure.

#### 2.1.5.5 All-*cis* 4-(2,3,6-trifluorocyclohexyl)benzoic acid (**60**)

Finally, and in order to complete the metabolism study, an incubation of all-*cis* 4-(2,3,6-trifluorocyclohexyl)benzoic acid (**60**) was carried out. Again, the benzylic hydroxylation product was isolated ( $t_{\text{ret}} = 23$  min). The reaction was relatively efficient, progressing to approximately 50% conversion in three days.



**Scheme 13.** *C. elegans* incubation with **60** gave **90**

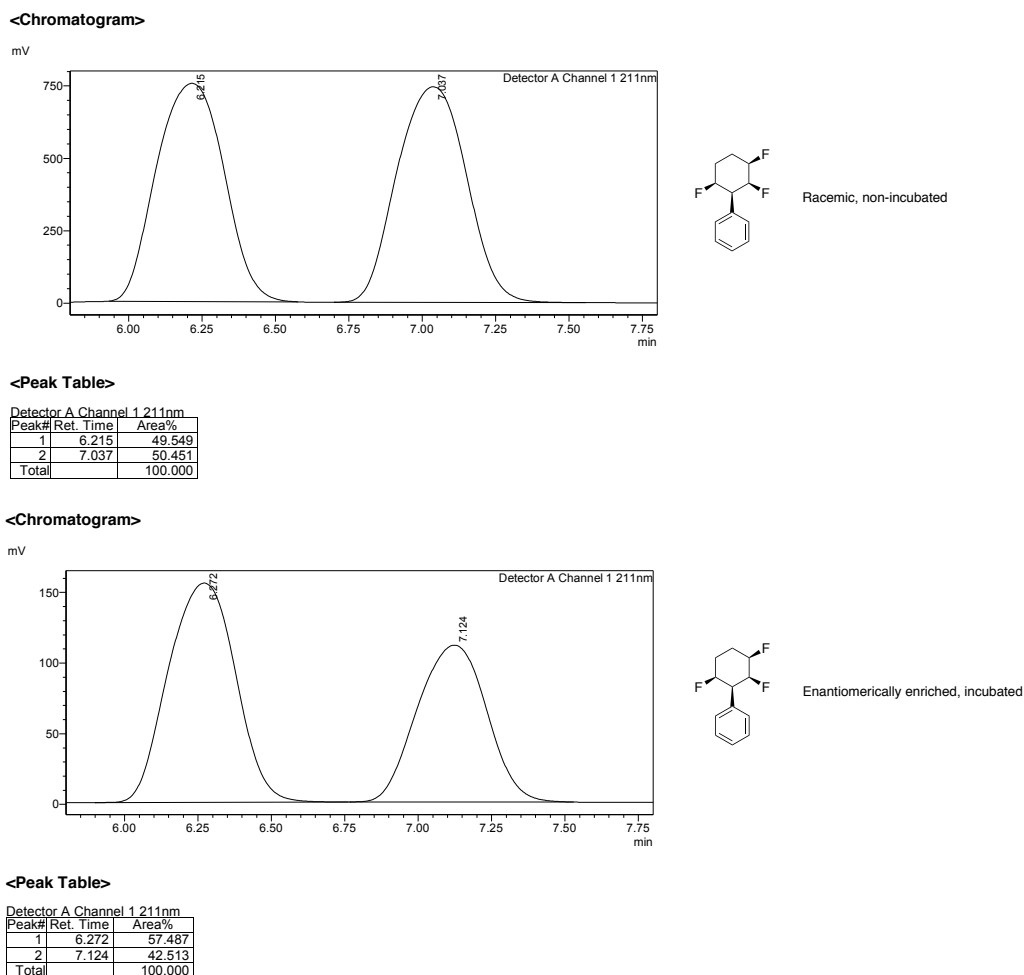
#### 2.1.6 Enantiomeric excess analysis of all-*cis* (fluorocyclohexyl)benzene derivatives

Compounds **58**, **59** and **60** fed to the fungus were all racemic mixtures. It became relevant to assess whether *C. elegans* mediated hydroxylations would be more selective towards one enantiomer over the other.

For chiral HPLC analysis, a sample of racemic all-*cis* (2,3,6-trifluorocyclohexyl)benzene **59** was injected, to establish the optimal separation conditions of the enantiomers. The sample recovered from the incubation was injected

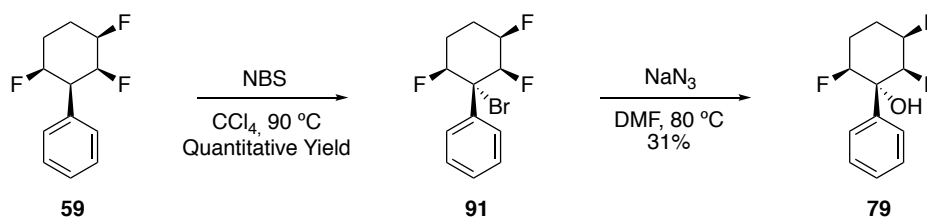


using the same conditions. HPLC enantiomeric analysis for racemic all-*cis* (2,3,6-trifluorocyclohexyl)benzene **59**, and after incubation with *C. elegans* (50% conversion) is shown in **Fig. 11**.



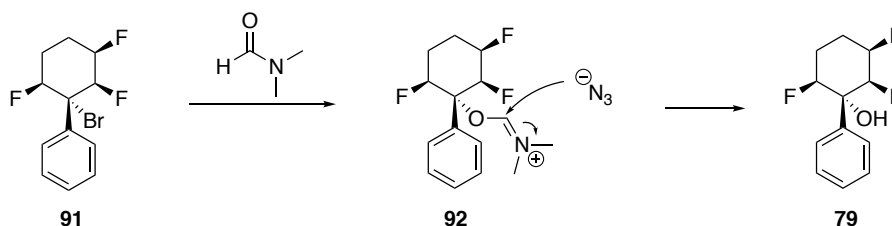
**Fig. 11** Chiral HPLC analysis for all-*cis* (2,3,6-trifluorocyclohexyl)benzene **59**. The upper trace is racemic **59**, and the lower trace is enantiomerically enriched (15% e.e.) after incubation with *C. elegans*

In order to assess the enantiomeric ratio of **79**, it became necessary to synthesise a racemic sample as a reference. The synthesis was carried out as shown in **Scheme 14**, and required a bromination step followed by introduction of the hydroxyl group by substitution.<sup>36</sup>



**Scheme 14.** Route to racemic 2,3,6-trifluoro-1-phenylcyclohexan-1-ol **79**

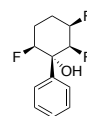
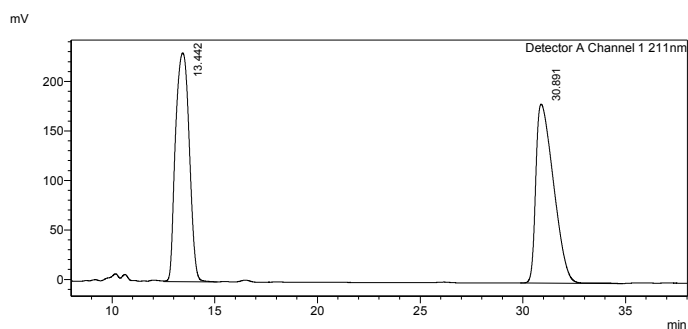
Benzylic bromination proved efficient, and allowed the required replacement of Br for OH. Following some recent work in the group, this last step was achieved in DMF to generate alcohol **79**.<sup>36</sup> A reasonable explanation for this process involves the participation of DMF as a nucleophile to form intermediate **92**, as recently published by T. Bykova in the group for the tetrafluorinated ring system.<sup>36</sup> This intermediate would then be attacked by the azide to release the benzylic alcohol, as shown in **Scheme 15**.



**Scheme 15.** Putative mechanism for the generation of alcohol **79**, where the oxygen source is DMF

Racemic 2,3,6-trifluoro-1-phenylcyclohexan-1-ol **79** was analysed by chiral HPLC, and a good separation of the enantiomers was achieved. A sample of 2,3,6-trifluoro-1-phenylcyclohexan-1-ol **79** produced from the incubation with *C. elegans* (50% conversion) was injected using the same conditions. This analysis (**Fig. 12**) indicates a 23% e.e.

<Chromatogram>

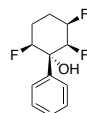
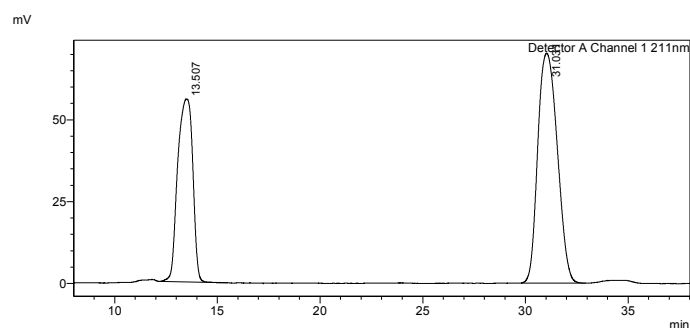


Racemic, non-incubated

<Peak Table>

Detector A Channel 1 211nm		
Peak#	Ret. Time	Area%
1	13.442	49.911
2	30.891	50.089
Total		100.000

<Chromatogram>



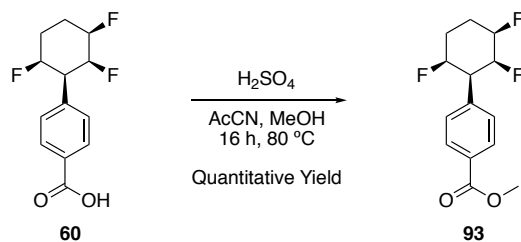
Enantiomerically enriched, incubated

<Peak Table>

Detector A Channel 1 211nm		
Peak#	Ret. Time	Area%
1	13.507	38.116
2	31.031	61.884
Total		100.000

**Fig. 12** Chiral HPLC analysis for 2,3,6-trifluoro-1-phenylcyclohexan-1-ol **79**. Upper trace shows the racemic mixture. Lower trace presents the enantiomerically enriched (23% e.e.) **79** obtained after incubation with *C. elegans* at 50% conversion

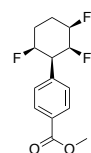
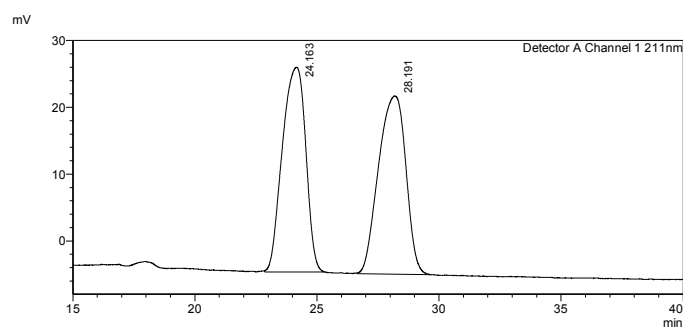
With these results in hand, the influence of the carboxylic acid group on enantiomeric selectivity was probed. However, it was not useful to inject carboxylic acid **60** directly onto the HPLC, as it is too polar, so it was derivatised as its methyl ester. The conditions and yield for this step are shown in **Scheme 16**, leading to quantitative formation of **93**.<sup>37</sup> Both the racemic and the metabolised samples were analysed by chiral HPLC, and led to identical racemic outcomes.



**Scheme 16.** Esterification conditions to **93**

The chiral HPLC results are presented in **Fig. 13**, for a 50% conversion in *C. elegans*.

<Chromatogram>

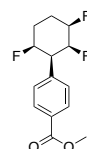
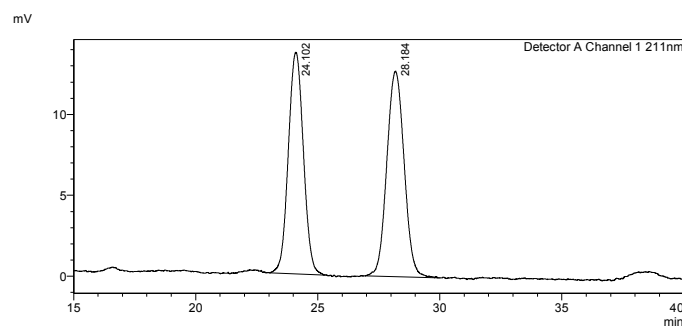


Racemic, non-incubated

<Peak Table>

Detector A Channel 1 211nm		
Peak#	Ret. Time	Area%
1	24.163	49.593
2	28.191	50.407
Total		100.000

<Chromatogram>



Enantiomerically enriched, incubated

<Peak Table>

Detector A Channel 1 211nm		
Peak#	Ret. Time	Area%
1	24.102	48.397
2	28.184	51.603
Total		100.000

**Fig. 13** Chiral HPLC analysis for **93**. The upper trace shows the racemic starting material and the lower trace corresponds to residual **93** after incubation with *C. elegans* (50% conversion)

### 2.1.7 Conclusions on the work with all-*cis* (fluorocyclohexyl) benzenes

*C. elegans* mimics human metabolism by performing Phase I modification by hydroxylation of xenobiotics. Lack of metabolism of an added xenobiotic would mean that the product is stable against metabolic degradation, and therefore, it would be unable to undergo Phase II metabolism and later be excreted from the organism. On the other hand, multiple hydroxylation, or a very active Phase I metabolism, would indicate poor metabolic stability, leading to low pharmaceutical efficiency.

In this section, the influence of the number of fluorine atoms on an otherwise identical cyclohexane motif was established. It was successfully shown that as the number of fluorine atoms increases, so does the metabolic stability, which is opposite to the effect of the fluorines on the lipophilicity. The estimations of lipophilicity by reverse-phase HPLC show that the Log *P* values decrease considerably as the number of fluorine atoms increases (and the compounds become more polar).

Both the lipophilicity and the metabolism studies are therefore promising for inclusion of the tetrafluoro motif in drug discovery programmes.

Enantiomeric excess analysis showed that the fungus is not particularly selective for the individual chiral compounds that were presented.

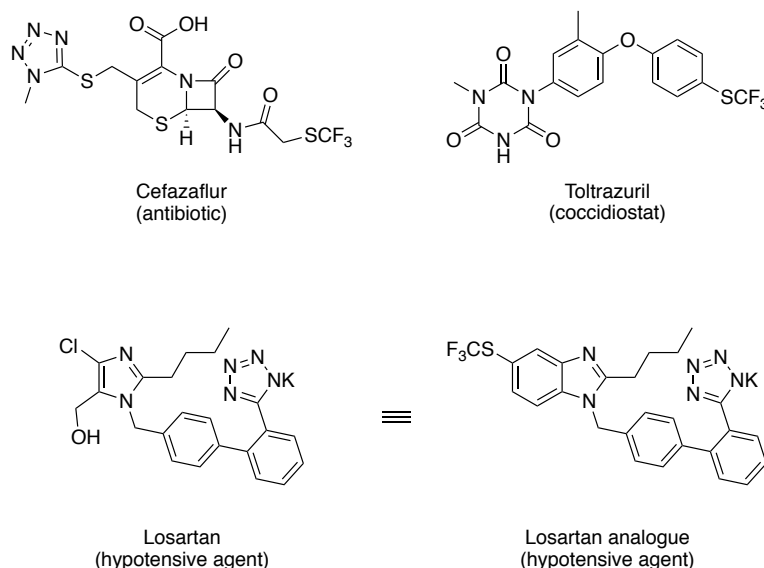
## 2.2 $\alpha,\alpha$ -Difluoroethyl thioethers and $\alpha,\alpha$ -difluoro oxyethers

### 2.2.1 Introduction to the syntheses and potential applications of $\alpha,\alpha$ -difluoroethyl thio- and oxy- ethers

This section focuses on  $\alpha,\alpha$ -difluoroethyl thioethers ( $R\text{-SCF}_2\text{CH}_3$ ).<sup>38</sup> This novel motif, introduced in St Andrews in 2018, has some similarity to both  $R\text{-XCF}_3$  and  $R\text{-XCH}_2\text{CF}_3$  ( $X = \text{S}$  or  $\text{O}$ ), moieties that are found quite widely in pharmaceutical drugs and agrochemicals.

Interest in trifluoromethyl sulfides as drug scaffolds derives from their high lipophilicity, strong electron-withdrawing properties and improved metabolic stability.<sup>39</sup> These characteristics improve the cell membrane permeability and biochemical stability of the parent molecules, making them ‘privileged structural motifs in drug discovery’.<sup>40</sup>

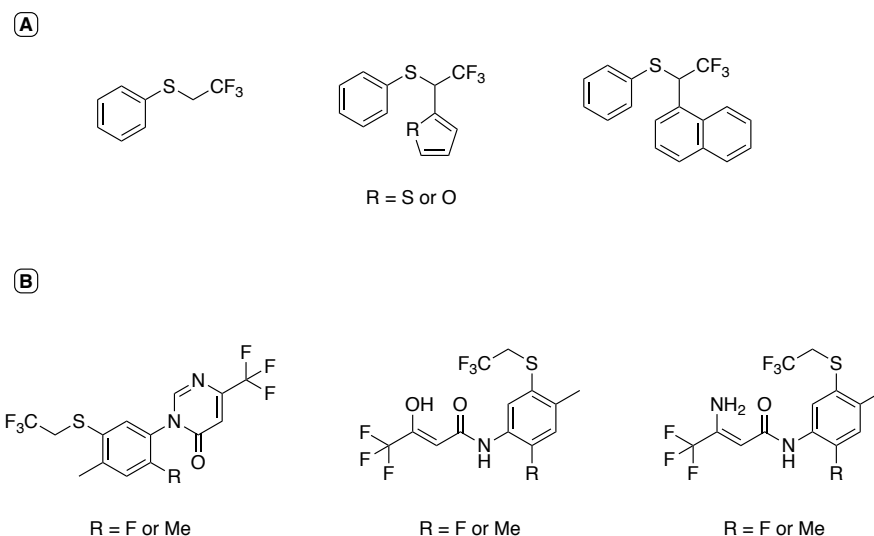
Some examples are shown in **Fig. 14**, and include cefazaflur (a first generation cephalosporin antibiotic), toltrazuril (an antiprotozoal veterinary agent) or a losartan analogue (a potential hypotensive drug).<sup>41, 42</sup>



**Fig. 14** Examples of drugs containing a  $\text{CF}_3\text{S-}$  motif

2,2,2-Trifluoroethyl sulfide derivatives are claimed quite frequently in patents. In 1990, for example, Uneyama patented the preparation of 2,2,2-trifluoro-1-

(phenylthio)ethanes for their use as pharmaceutical and agrochemical intermediates (**Fig. 15a**).<sup>43</sup> More recently, Koehler and coworkers reported the preparation of several trifluoroethyl sulfide derivatives for pest control (**Fig. 15b**).<sup>44</sup>



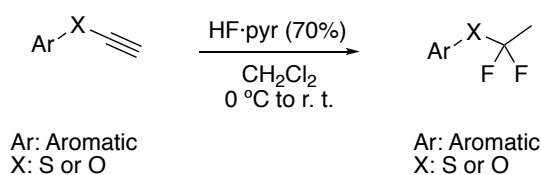
**Fig. 15** Examples of trifluoroethyl sulfides in patents

Analyses of the drug Riluzole, and its sulfuric analogue, SKA-19 (**Fig. 16**) were also performed by the St Andrews group.<sup>38</sup> Riluzole is used to treat amyotrophic lateral sclerosis, as well as anxiety,<sup>38,45</sup> while its sulfur containing analogue SKA-19 is a very potent anticonvulsant.<sup>38,46</sup>



**Fig. 16** Structures of riluzole and SKA-19, along with the expected bioisosteres

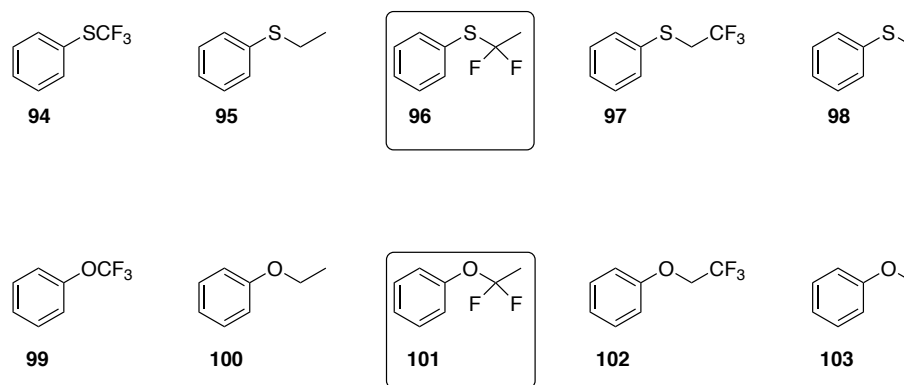
The synthetic method used to obtain these  $\alpha,\alpha$ -difluoroethyl ethers is shown in **Scheme 17**, and involves the double hydrofluorination of the corresponding ethynyl ethers with HF·pyr (70%).<sup>38</sup>



**Scheme 17.** General procedure for the generation of selected difluorinated thio- and oxy-ethers<sup>38</sup>

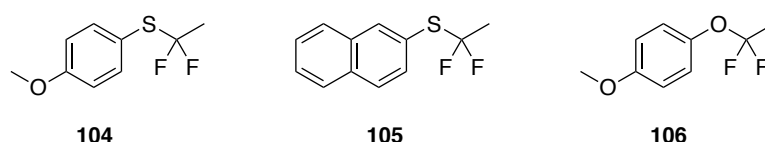
### 2.2.2 Aims and Objectives

It became an objective to assess the lipophilicity profile ( $\log P$ ) of selected  $\alpha,\alpha$ -difluoroethyl thio and oxy ethers (**96** and **101**), relative to reference compounds (**Fig. 17**), given the very recent introduction of this motif.



**Fig. 17** Selected  $\alpha,\alpha$ -difluoroethyl thio and oxy ethers for the studies on lipophilicity, and the chosen compounds for comparison. Compound **96** was synthesised by Dr. Tomita (University of St. Andrews), and compound **101** was prepared for this project. The rest of the products were commercially available

Additionally, the metabolic stability of the difluoro motif – and therefore its metabolic fate – was investigated to obtain preliminary data that could inform future drug discovery programmes. The selected molecules are shown below (**Fig. 18**).



**Fig. 18** Selected  $\alpha,\alpha$ -difluoroethyl thio and oxy ethers for the studies on metabolism. Products **104** and **105** were synthesised by Dr. Tomita and **106** was obtained by Dr. Al-Maharik (University of St. Andrews)

### 2.2.3 Estimation of lipophilicities by the use of a reverse-phase HPLC method

Estimation of the lipophilicity was carried out using a similar reverse-phase HPLC procedure similar to that used for the fluorinated cyclohexyl derivatives, without TFA added.

The set of reference compounds used was the optimised range as described before, and shown in **Table 6**, with retention times established for the non-TFA added eluent

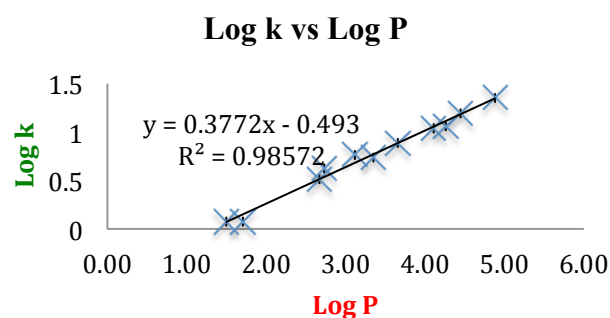


conditions. As expected, the changes in retention times are not significant, but re-calibration was required to ensure accuracy.

**Table 6.** Log *P* values for the compounds used to establish a reference curve

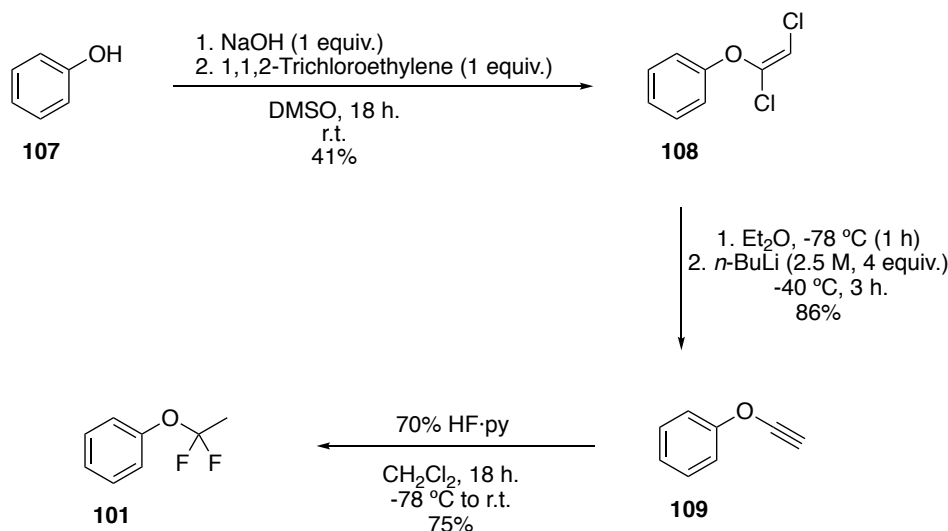
Reference	Log <i>P</i>	Rt 1 (min)	Rt 2 (min)	Rt 3 (min)	Average Rt (min)	Capacity factor (k)	Log k
Phenol	1.50	4.30	4.30	4.26	4.29 ± 0.03	1.18	0.08
2-Fluorophenol	1.71	4.20	4.21	4.50	4.30 ± 0.20	1.18	0.08
Benzofuran	2.67	8.38	8.38	8.38	8.38 ± 0.00	3.26	0.51
Toluene	2.73	10.26	10.27	10.25	10.26 ± 0.01	4.22	0.63
<i>o</i> -Xylene	3.12	13.25	13.28	13.58	13.37 ± 0.30	5.80	0.76
Naphthalene	3.35	12.73	12.73	12.76	12.74 ± 0.02	5.48	0.74
Cumene	3.66	17.06	17.07	17.17	17.10 ± 0.07	7.70	0.89
<i>t</i> -Butylbenzene	4.11	23.45	23.41	23.26	23.37 ± 0.11	10.89	1.04
Butylbenzene	4.26	29.96	30.43	30.45	30.28 ± 0.32	14.41	1.16
Anthracene	4.45	24.46	24.46	24.51	24.48 ± 0.03	11.46	1.06
Pyrene	4.88	32.70	32.62	32.80	32.71 ± 0.09	15.65	1.19

The values in **Table 6** were used to establish a new regression curve, by plotting the calculated Log k against the reported Log *P* (**Plot 2**).



**Plot 2.** Log k vs. Log *P* for the new neutral conditions, with no supplementation of TFA

This regression line was used to calculate the Log *P* values for the thio- and oxyethers series, as presented in **Table 7**. For the estimation of **101**, it was necessary to synthesise this compound, as illustrated in **Scheme 18**.<sup>38</sup>

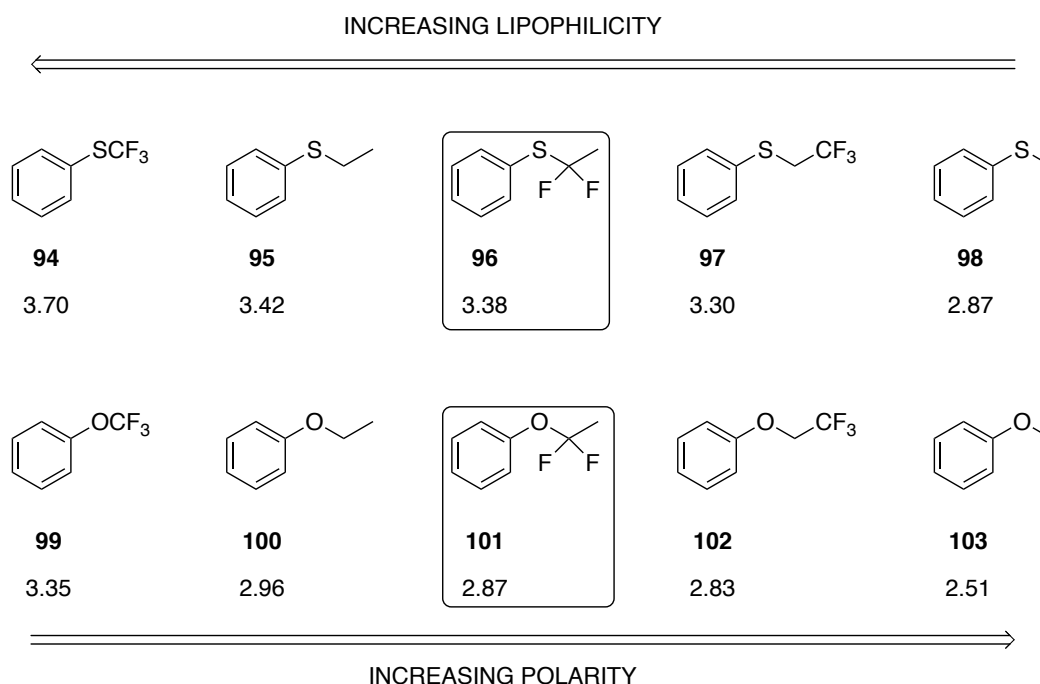


**Scheme 18.** Synthetic procedure for the generation of **101**

**Table 7.** Measured lipophilicity values for a series of fluorinated thio- and oxy ethers

Compound	Log <i>P</i>	Rt 1	Rt 2	Rt 3	Average Rt	Capacity factor ( <i>k</i> )	Log <i>k</i>
PhSCF <sub>3</sub> ( <b>94</b> )	3.70	17.67	17.68	17.68	17.68 ± 0.01	8.00	0.90
PhSCH <sub>2</sub> CH <sub>3</sub> ( <b>95</b> )	3.42	14.32	14.28	14.33	14.31 ± 0.03	6.28	0.80
PhSCF <sub>2</sub> CH <sub>3</sub> ( <b>96</b> )	3.38	13.89	13.88	13.88	13.88 ± 0.01	6.07	0.78
PhSCH <sub>2</sub> CF <sub>3</sub> ( <b>97</b> )	3.30	13.05	13.05	13.03	13.04 ± 0.01	5.64	0.75
PhSCH <sub>3</sub> ( <b>98</b> )	2.87	9.61	9.61	9.62	9.61 ± 0.01	3.89	0.59
PhOCF <sub>3</sub> ( <b>99</b> )	3.35	13.49	13.51	13.49	13.50 ± 0.01	5.87	0.77
PhOCH <sub>2</sub> CH <sub>3</sub> ( <b>100</b> )	2.96	10.42	10.09	10.12	10.21 ± 0.21	4.20	0.62
PhOCF <sub>2</sub> CH <sub>3</sub> ( <b>101</b> )	2.87	9.56	9.57	9.58	9.57 ± 0.01	3.87	0.59
PhOCH <sub>2</sub> CF <sub>3</sub> ( <b>102</b> )	2.83	9.28	9.28	9.48	9.35 ± 0.13	3.76	0.58
PhOCH <sub>3</sub> ( <b>103</b> )	2.51	7.55	7.60	7.57	7.57 ± 0.03	2.85	0.46

These results are illustrated in **Fig. 19**, in a different format to **Table 7**, where a comparison is made between the lipophilicity and polarity trends.<sup>38</sup>



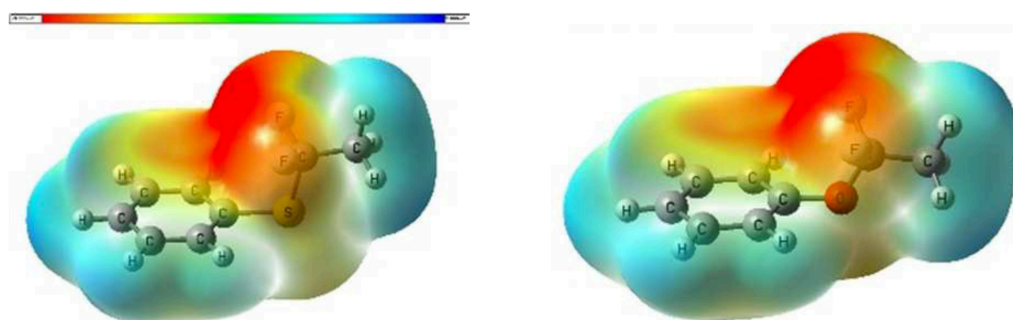
**Fig. 19** Lipophilicity versus polarity for a series of thio- and oxy- ethers<sup>38</sup>

The trends in lipophilicity and polarity observed during this investigation are informative. It is well-known that lipophilicity decreases upon monofluorination, but the trend reverses with vicinal trifluorination.<sup>47</sup> Difluorination is generally similar to the values of monofluorination.<sup>1,2,3,47</sup> In our series of thio- and oxy- ethers, the highest Log *P* values are observed with trifluoromethoxy- (**99**) and trifluorothio- (**94**) anisole; which contrast with the lowest Log *P* values of their hydrocarbon analogues anisole **103** and thioanisole **98**.<sup>38</sup> For the compounds **95**, **96** and **97**; as well as their oxy derivatives, intermediate polarities are found.

Partial fluorination generates orientated dipoles.<sup>47</sup> Fluorine atoms also have a marked inductive effect, which results in the polarisation of their vicinal hydrogen atoms.<sup>38</sup> Studies on geminal  $\alpha,\alpha$ -difluoropropyl benzenes have concluded that the volume increase produced by fluorine substitution increases lipophilicity. In this way, although polarity is increased by the addition of a second fluorine atom, it is compensated for by the subsequent volume increase.<sup>47a</sup> Overall, this buffers the inductive effect and leads to an unchanged lipophilicity value with respect to the monofluorinated analogue,

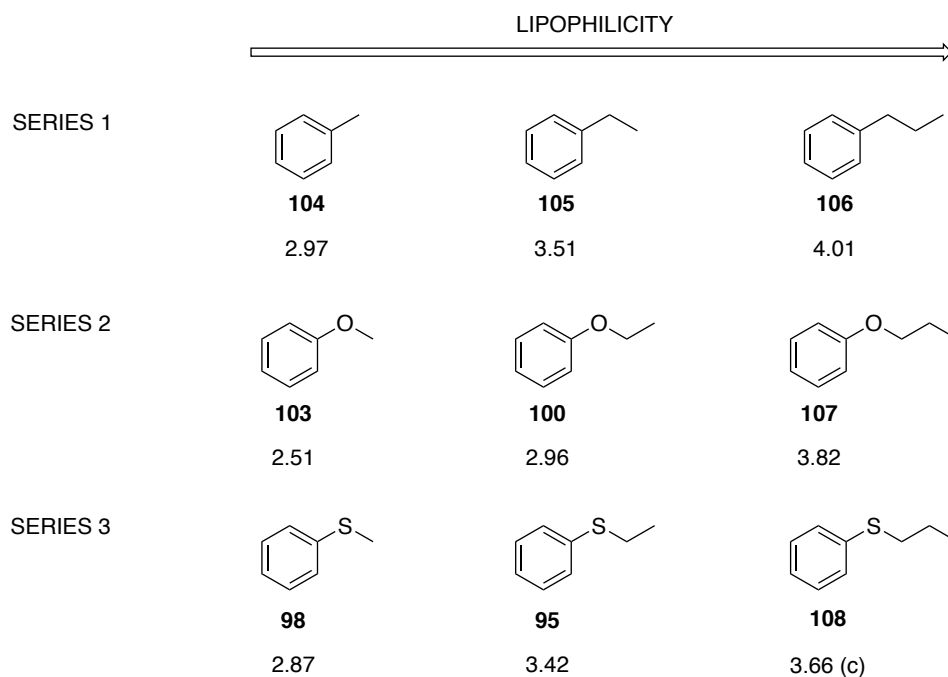
although lower than the non-fluorinated derivative.<sup>47a</sup> These factors help explain the observed Log *P* values for the  $\alpha,\alpha$ -difluoroethyl thio (**96**) and oxy (**101**) ethers, and their non-fluorinated analogues (**95** and **100**). Additionally, these new motifs have a bridging sulfur or oxygen atom, both of which possess lone pairs. However, the negative inductive effect of the fluorine atoms will suppress lone pair donor capacity, and influence the lipophilicity values.

**Fig. 20** shows the electrostatic surface potential maps for the thio ether (left) and the oxygen ether (right).<sup>38</sup>



**Fig. 20** Electrostatic surface potential maps for the thio- and oxy- analogues **96** and **101**. Computational analysis and calculations were performed by M. Bühl (University of St Andrews)<sup>38</sup>

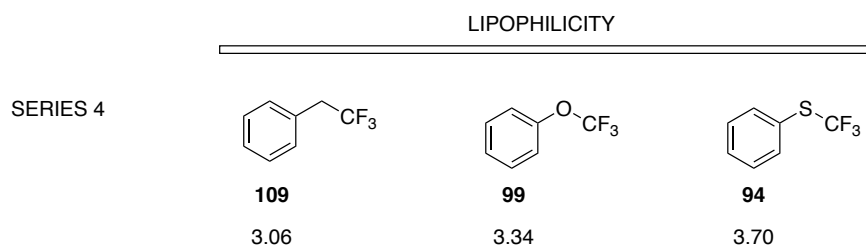
Additionally, some other products were analysed in this lipophilicity study. For an easier reading and analysis of the results, the compounds have been organised in series 1 to 6, in **Figs. 21-23**.



**Fig. 21** Series 1 to 3. Log *P* values comparing the alkyl and ether chain lengths. (c) in **108** refers to a calculated Log *P*

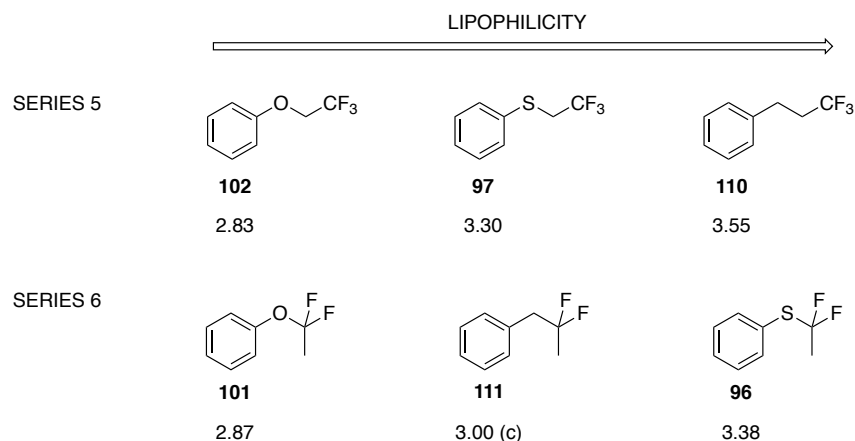
Series 1 reflects how the carbon chain length influences lipophilicity. The comparison of toluene, ethylbenzene and propylbenzene illustrates that the longer the carbon chain, the higher the Log *P* value. The same trends are observed for the oxygen and thio ethers in series 2 and 3. If these three series are compared, the behavioural trends are common, although the lipophilicity values are dissimilar. The absolute values on Series 1 are the highest, followed by Series 3. The oxygen ethers of Series 2 display the lowest lipophilicity values of the three.

Series 4 (**Fig. 22**) highlights the effect of trifluorination for 2-membered chains attached to phenyl. Therefore, trifluoroethylbenzene **109**, trifluoroanisole **99** and trifluorothioanisole **94** are compared. Interestingly, the non-ether **109** is the most polar. Series 5 and 6 are analogous to Series 4, for different length chains and different fluorinated motifs.



**Fig. 22** Series 4. Log *P* comparisons of two-membered side chains with CF<sub>3</sub> groups

Series 4 shows that trifluoroethylbenzene **109** has the lowest Log *P* value, followed by the oxy and thio ether derivatives, **99** and **94** respectively. This highest lipophilicity is presumably due to the larger volume of sulfur, which actively softens the nucleophilicity of the lone pairs, reducing hydrogen bonding capacity and therefore resulting in a higher lipophilicity value compared to the oxygen analogue. The presence of the oxygen atom close to fluorine will decrease overall lipophilicity relative to the parent compound.<sup>48</sup> In this case, this trend is not observed, and trifluoroethyl benzene **109** presents a lower Log *P* value than trifluoromethoxy benzene **99**. This perhaps unexpected result can be explained by fluorine polarising the vicinal C-H bonds.

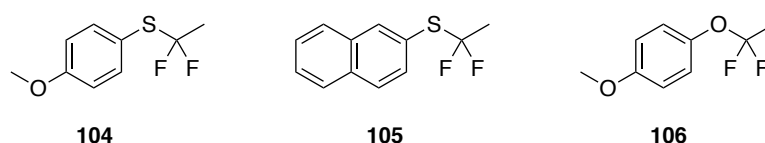


**Fig. 23** Log *P* comparison of three membered side chains containing PhXCH<sub>2</sub>CF<sub>3</sub> and PhXCF<sub>2</sub>CH<sub>3</sub> motifs (X= C, O or S)

Series 5 and 6 follow this behaviour too, when the sulfur and oxygen structures are compared. In this way, the oxy ethers emerge with the lowest lipophilicities. Interestingly, when the Ph-XCH<sub>2</sub>CF<sub>3</sub> and the Ph-XCF<sub>2</sub>CH<sub>3</sub> motifs (X= S or O) in these series are compared, the lipophilicity values are quite similar, indicating that they are good isosteres for one another.

## 2.2.4 Metabolic fate of the ethers in *Cunninghamella elegans*

In order to explore the metabolic fate of these ether derivatives, three candidates were selected, based on their availability and low volatility. Accordingly, the selected compounds were (1,1-difluoroethyl)(4-methoxyphenyl)sulfane (**104**), (1,1-difluoroethyl) (naphthalene-2-yl)sulfane (**105**) and 1-(1,1-difluoroethoxy)-4-methoxybenzene (**106**). These compounds, shown in **Fig. 24**, were prepared in the University of St Andrews by Dr Tomita (**104** and **105**) and Dr Al-Maharik (**106**). Incubation, isolation and purification procedures followed those previously described.



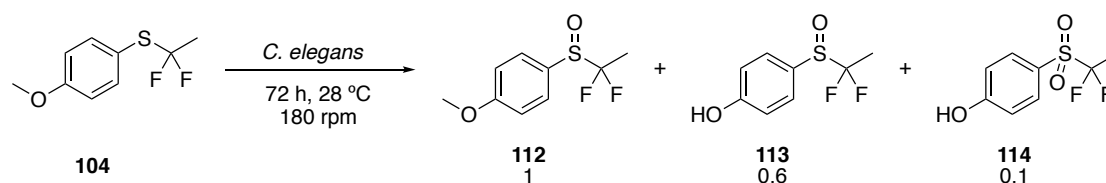
**Fig. 24** Selected molecules for the study on metabolic fate in *C. elegans*

### 2.2.4.1 (1,1-Difluoroethyl)(4-methoxyphenyl)sulfane (**104**)

Incubation of thioether **104** in *C. elegans* led to the isolation of three new fluorometabolites, products of oxidation from the fungal CYP enzymes.

Analysis of the extracts by <sup>19</sup>F{<sup>1</sup>H}NMR spectroscopy showed a clear feature. While the starting material appears as a singlet peak for the fluorines, two of the fluorometabolites displayed a doublet of doublets, indicating an AB system and a structural change in the close environment of the fluorine atoms.

Isolation of each peak by reverse-phase HPLC, and subsequent NMR spectroscopy and high-resolution mass spectrometry, revealed sulfoxides **112** and **113** and sulfone **114**.



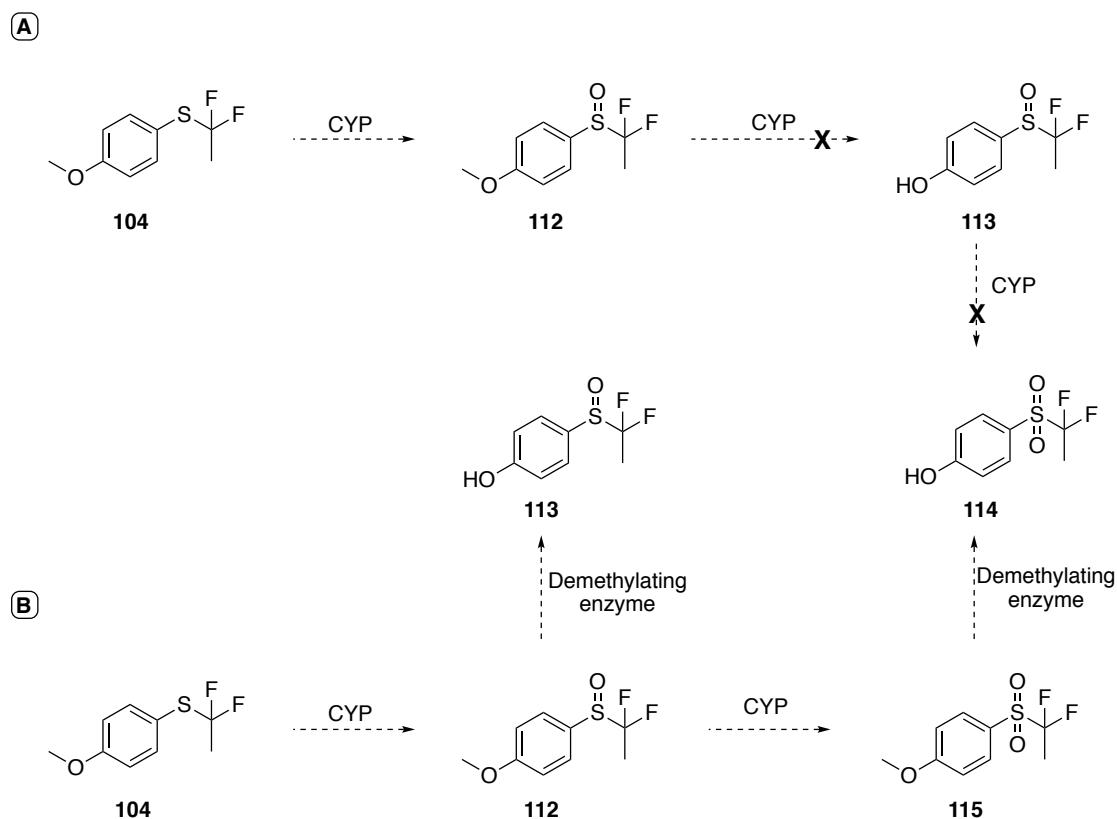
**Scheme 19.** Fluorometabolites isolated, along with their isolation ratios

The major product was sulfoxide **112** ( $t_r$  = 24 min), followed by phenol **113** ( $t_r$  = 16 min). Sulfone **114** ( $t_r$  = 21 min) was a relatively minor metabolite, at only 10% of

sulfoxide **112**. The sulfoxides are chiral on the sulfur atom, resulting in non-equivalence of the fluorine atoms, and hence, the observed AB system. The lower level of sulfone **114** could be an indication that it is the last product to be generated in the metabolic pathway.

Additional to sulfur oxidation, *C. elegans* is also able to de-methylate the *para* methoxyl group, generating a phenol. The fact that there is no oxidation of the aromatic ring, nor elimination of the fluorines, seems to point towards a reasonably stable motif overall.

In an effort to establish the sequence of events through which these metabolites are generated, a series of control experiments was carried out. It was hypothesised that sulfoxide **112** would be generated first. A demethylation would then act to generate **113**, which would finally be oxidised to sulfone **114** (**Scheme 20a**). Alternatively, the final steps could possibly invert in order, and generate sulfone **115** prior to demethylation (**Scheme 20b**).



**Scheme 20.** Hypothetical pathways for (1,1-difluoroethyl)(4-methoxyphenyl)sulfane (**100**) metabolism



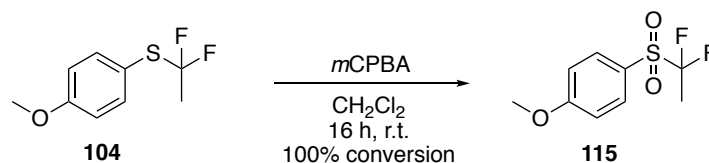
In order to establish if any of these routes is valid, each of the fluorometabolites was separately incubated, with different rates of success. It proved difficult to establish outcomes by  $^{19}\text{F}\{^1\text{H}\}$  NMR spectroscopy from the crude extracts, so it was necessary to use HPLC. Comparison of retention times and NMR analyses of the collected peaks proved more successful and allowed characterisation.

Incubation of sulfoxide **112** ( $t_r = 24$  min) led to the formation of phenol **113** and sulfone **114**, both being securely identified against previous data. No starting material (**112**) was observed either by NMR spectroscopy or HPLC, indicating full conversion. These results prove that oxidation of the sulfide happens first in the metabolism.

An incubation of sulfoxide **113** ( $t_r = 16$  min) with *C. elegans* did not result in any conversion and only the starting material **113** was recovered.

Finally, given the very low amount that was recovered from the original culture, an experiment with sulfone **114** did not show any new products, indicating that the metabolite is either exuded and not taken up by the fungus, or possibly fully metabolised.

These experiments established the first step of the metabolism, and also allowed route **A** of **Scheme 21** to be discarded. In an attempt to establish whether route **B** is valid, the sulfone **115** was synthesised for an incubation study (**Scheme 21**). Given that the synthesis was carried out on analytical amounts of **104** (5 mg), a complete conversion was estimated by NMR spectroscopy only, but the yield could not be established.



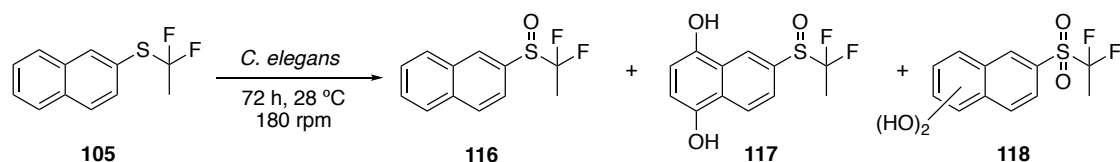
**Scheme 21.** Synthesis of 1-((1,1-difluoroethyl)sulfonyl)-4-methoxybenzene (sulfone **115**)

Sulfone **115** was incubated with *C. elegans*, following the standard conditions. After 72 h, the culture was extracted and the crude analysed.  $^{19}\text{F}\{^1\text{H}\}$  NMR spectroscopy showed two peaks (singlets), which match the starting material **115** and sulfone **114**. This is consistent with the last step in **Route B**, and gives a stronger support for **Route B** over **Route A**.

#### 2.2.4.2 (1,1-Difluoroethyl)(naphthalene-2-yl) sulfane (**105**)

In order to study the influence of the *p*-methoxy substituent on the metabolism of the thioether motif, a derivative of difluorinated **104** without the –OMe group was explored. (1,1-Difluoroethylphenyl)sulfane would have been the most obvious choice, but this compound is too volatile for the complete cycle of incubation, extraction and purification. Instead, the metabolic fate of (1,1-difluoroethyl) (naphthalene-2-yl) sulfane (**105**) was investigated.

After incubation with *C. elegans*,  $^{19}\text{F}\{^1\text{H}\}$  NMR analysis of the extract showed three sets of signals, indicating the formation of two sulfoxide metabolites and one sulfone (**Scheme 22**) – reproducing the observed products in the methoxy derivative **104**.



**Scheme 22.** *C. elegans* incubation results for **105**

Reverse-phase HPLC allowed the separation of these metabolites. NMR analysis, followed by high-resolution mass spectrometry enabled a preliminary identification of each of the structures. However, the recovery of fluorometabolites from the fungal culture was low, making full characterisation challenging.

The major product ( $t_r = 37$  min) was positively identified as sulfoxide **116**. The second most abundant fluorometabolite ( $t_r = 13$  min) is sulfoxide **117**, as determined by  $^1\text{H}$  and  $^{19}\text{F}\{^1\text{H}\}$  NMR spectroscopy, as well as high-resolution mass spectrometry. Unfortunately, the amount of material recovered was too low to be able to carry out  $^{13}\text{C}$  NMR analysis. The  $^1\text{H}$  NMR spectrum is however quite distinctive, and points towards this structure. In particular, three signals in the aromatic region, between 7.27 and 7.75 ppm, match the substitution pattern for the aromatic ring adjacent to the sulfur atom; while two doublets at 6.5 and 6.1 ppm correspond to protons vicinal to the hydroxyl groups.

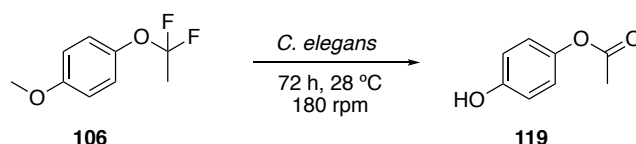
Finally, a third fluorometabolite of general sulfone structure **118** ( $t_r = 17$  min), according to mass spectrometry, was recovered by HPLC. This is a very minor metabolite and little information was obtained for its unambiguous identification.

#### 2.2.4.3 1-(1,1-Difluoroethoxy)-4-methoxybenzene (**106**)

Finally, in order to extend the study to the oxy ether series, difluoroethoxy ether **106** (5 mg) was fed to cultures of *C. elegans*, employing the standard conditions.

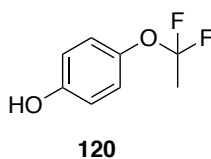
Examination of the  $^{19}\text{F}$  NMR spectrum of the extracts showed only residual starting material, although only at low levels. Isolation of the major components by HPLC and their subsequent analysis proved that there was a non-fluorinated metabolite present.

This was determined to be 4-acetoxyphenol **119** ( $t_r = 25$  min), the identity of which was confirmed by comparison of literature NMR data.<sup>49</sup>



**Scheme 23.** Incubation of **106** in *C. elegans*, gave 4-acetoxyphenol **119**

These results were reproduced in separate experiments (4 times). However, on one occasion, the  $^{19}\text{F}$  NMR spectrum showed the formation of a new fluorine peak. HPLC purification showed that the main metabolite was still 4-acetoxyphenol **119**, but also allowed the isolation of this new fluorinated compound ( $t_r = 19$  min). Unfortunately, the recovery from the fungal culture was once again very low, and taking into account that this was a very minor metabolite, only  $^1\text{H}$  and  $^{19}\text{F}$  NMR spectra were obtained. These analyses were enough to observe that the  $\text{OCF}_2\text{CH}_3$  motif was intact, and that demethylation had occurred. A residual *para* substitution pattern was obvious in the  $^1\text{H}$  NMR spectrum, suggesting 4-(1,1-difluoro)ethoxy phenol **120**.



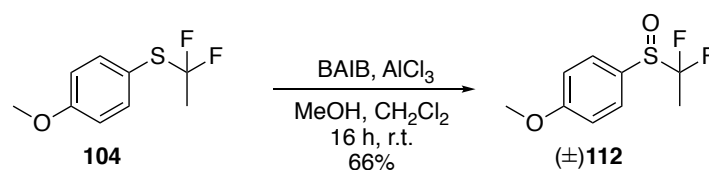
In order to further establish what happened in this incubation, scale-up ( $3 \times 25$  mg) batches of **106** were added to three different cultures of *C. elegans*. The outcome of each was exactly the same as that for the first assays, and showed the formation of 4-acetoxyphenol **119** along with some residual starting material, and some minor traces of decomposition, but no phenol **120**.

Therefore, it was determined that 1-(1,1-difluoroethoxy)-4-methoxybenzene (**106**) suffers from a rapid hydrolytic metabolism, as well as demethylation, and seems to be a vulnerable motif for bioactive discovery programmes, given its very low metabolic stability.

### 2.2.5 Enantiomeric excess analysis

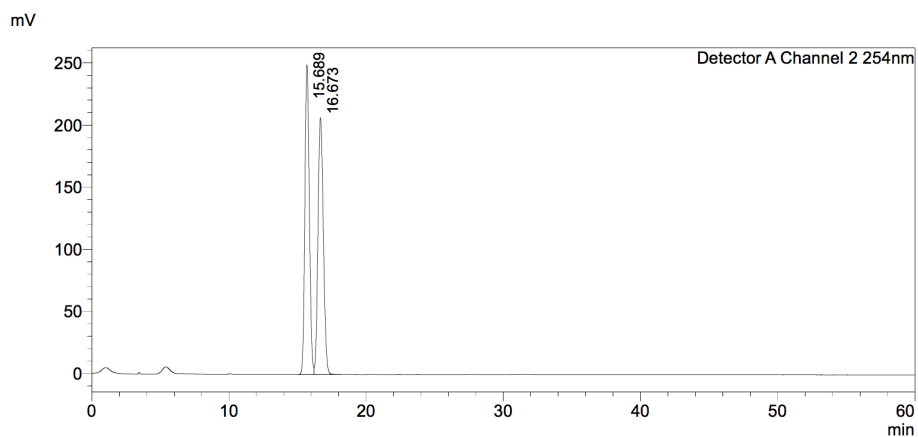
The experiments with thioether derivatives (**104** and **105**) led to the isolation of chiral sulfoxide intermediates. The enantiomeric selectivity of the fungus in generating these chiral sulfoxide motifs was also investigated.

Initially, a racemic mixture of 1-((1,1-difluoroethyl)sulfinyl)-4-methoxybenzene (**112**) was synthesised by addition of (bisacetoxy)iiodo benzene (BAIB) and aluminium chloride to (1,1-difluoroethyl)(4-methoxyphenyl)sulfane (**104**), as illustrated in **Scheme 24**. This reaction was carried out on an analytical scale (4.7 mg), and a portion of the product was purified by reverse-phase HPLC.



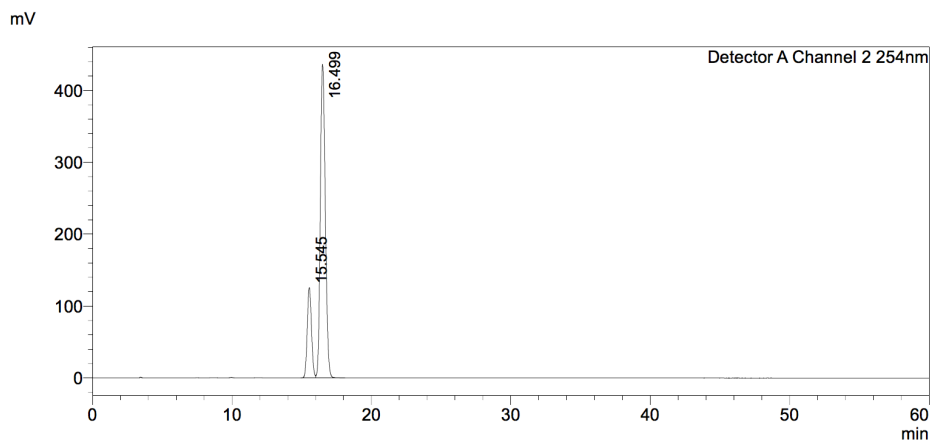
**Scheme 24.** Synthesis of racemic **112**

A solution of racemic sulfoxide **112** in isopropanol was analysed by chiral HPLC. The successful eluent conditions afforded a separation of 1 min between the enantiomers, and involved the use of an ID column (5% isopropanol in hexane, 1 mL per minute). The same conditions were used to analyse the product recovered from the fungal culture after incubation with **104**. In this case, the fungus manages to generate sulfoxide **112** in 60% e.e (**Fig. 25**).



<Peak Table>

Detector A Channel 2 254nm		
Peak#	Ret. Time	Area%
1	15.689	49.906
2	16.673	50.094
Total		100.000

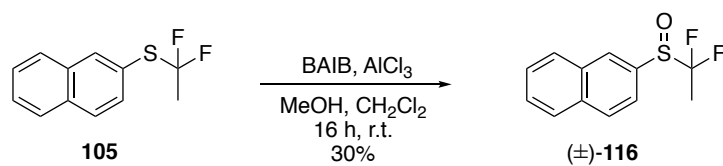


<Peak Table>

Detector A Channel 2 254nm		
Peak#	Ret. Time	Area%
1	15.545	19.640
2	16.499	80.360
Total		100.000

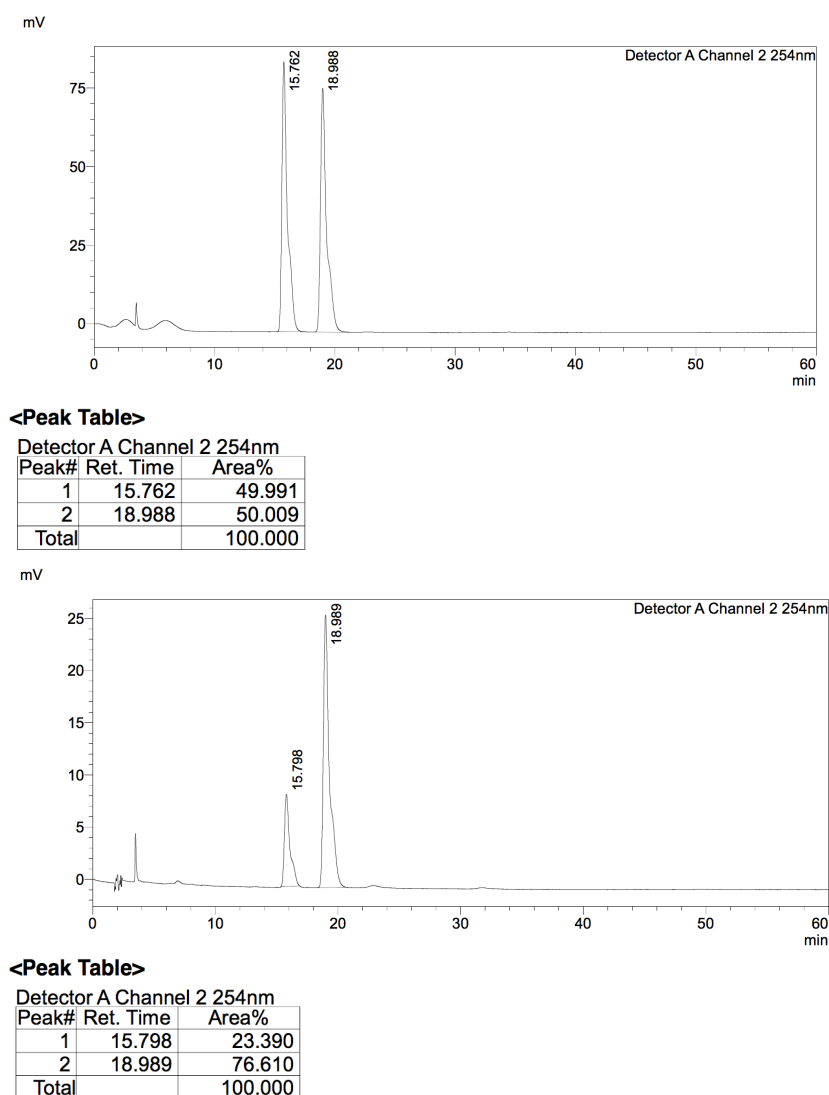
**Fig. 25** Chiral HPLC traces for the racemic **112** (upper image) and the isolated enantiomerically-enriched **112** after incubation of **104** with *C. elegans* (lower image)

The same synthetic procedure was used to obtain racemic sulfoxide **116** on an analytical scale (**Scheme 25**).



**Scheme 25.** Synthesis of racemic **116**

The fully characterised racemate was successfully separated into its enantiomers by chiral HPLC (IC column with 5% isopropanol in hexane at a rate of 1 mL/min). Analysis of the fungal extracts indicated that, for this naphthalene derivative, the fungus performs the biotransformation with a 54% e.e. Although this value is slightly lower than that of the methoxy derivative, it is still quite selective (**Fig. 26**).



**Fig. 26** Chiral HPLC traces for racemic mixture **116** (upper trace) and the isolated enantiomerically-enriched **116** after incubation of *C. elegans* with **105** (lower trace)

### 2.2.6 Conclusions on $\alpha,\alpha$ -difluoroethyl thio- and oxy- ethers

The lipophilicity profile of a range of  $\alpha,\alpha$ -difluoroethyl thioethers and  $\alpha,\alpha$ -difluoro oxyethers was estimated by reverse-phase HPLC.

In this way, the metabolic fate of selected examples was carried out by incubations with *C. elegans*. While the sulfur thio ethers display a moderate metabolic stability, yielding sulfoxide and sulfone metabolites, the investigated oxy ether was readily metabolised to an acetate ester, indicating low metabolic stability.

## 2.3 $\alpha,\beta,\beta$ -(Trifluorocyclopropyl) benzene derivatives

Cycloalkanes are present in countless natural and synthetic products, all of which display an ample variety of applications, from antibiotics to liquid crystals – and many more examples in-between. For a number of years, the St Andrews group has been studying the synthesis and properties of all-*cis* fluorocyclohexanes. Recently, the group has extended this work to the synthesis of partially fluorinated cyclopropanes for their potential in drug discovery programmes.

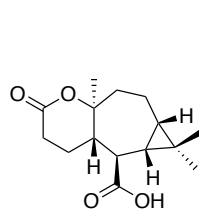
### 2.3.1 Introduction to the syntheses and potential applications of fluorocyclopropanes

The first cyclopropyl derivative reported was synthesised by W. H. Perkin in 1884.<sup>50</sup> Since then, the literature on cyclopropane syntheses has been continuous.<sup>51</sup> The versatility of this moiety is reflected in its multiple molecular purposes; they can act as ‘space elements’ of a specific dimension, reduce conformational flexibility, store strain energy to produce a thermodynamic driving force, or serve as high-energy intermediates in metabolism, among others.<sup>52</sup> In 2003, there were over 200 pharmaceutically relevant drugs that contained at least one cyclopropylamine moiety.<sup>53</sup>

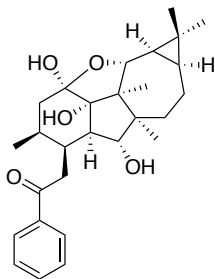
The presence of cyclopropyl rings in natural products is also significant, and they have been isolated from different plants, fungus and both marine and terrestrial microorganisms.

**Fig. 27** shows a selection of natural products containing, at least, one cyclopropane moiety.<sup>54</sup> These products display a variety of biological activities,<sup>55, 56</sup> from antiviral,<sup>57</sup> antifungal<sup>58, 59, 60</sup> and antibacterial<sup>61</sup> to anticancer<sup>62</sup> and anti-leishmanial.<sup>63</sup>

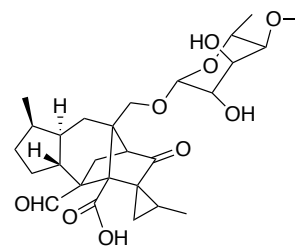




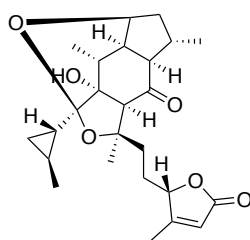
**Volvalerelactone A**  
(Unknown activity)



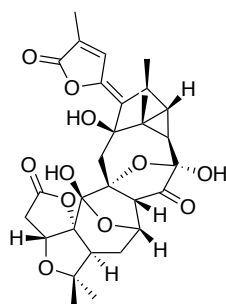
**Euphorbactin**  
(Activity against HIV-1 replication  
and antiviral activity)



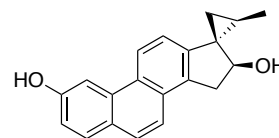
**Sordarin E**  
(Antifungal activity)



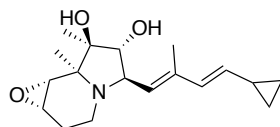
**Colquhounoid A**  
(Antifungal activity)



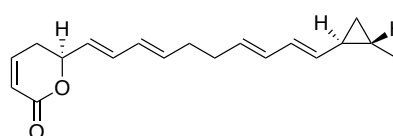
**Lancolide A**  
(Antiplatelet aggregation)



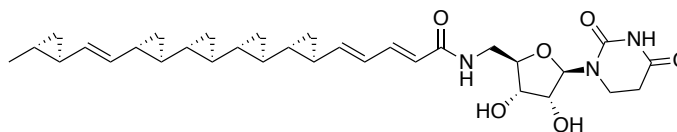
**Cinanthrenol A**  
(Cytotoxicity against P-388  
and HeLa cancer lines)



**Cyclizidine**  
(Antibiotic activity)



**Coibacin A**  
(Anti-leishmanial and  
antiinflammatory activity)

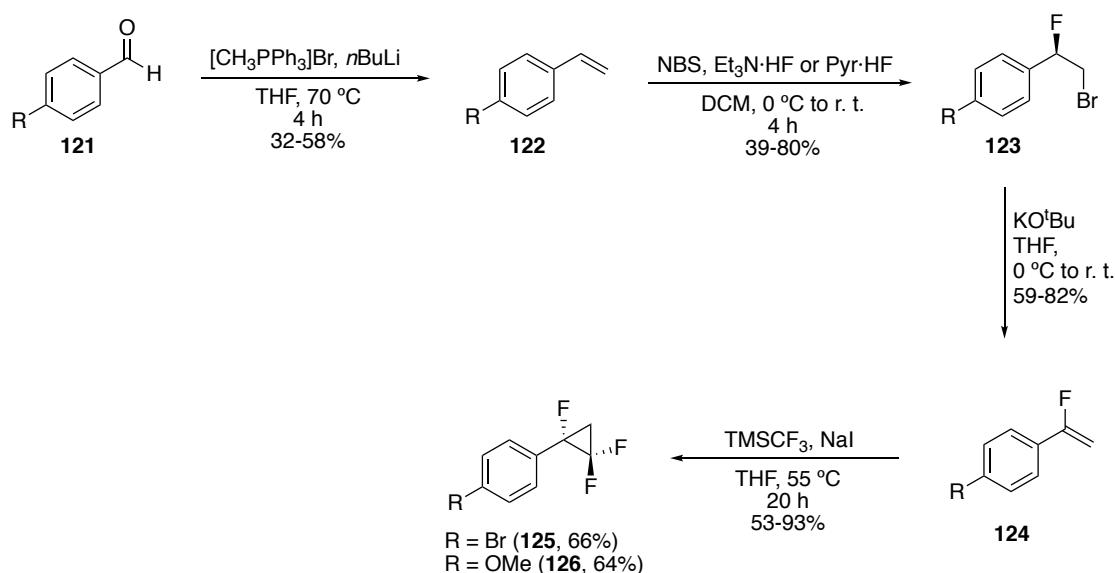


**Jawsamycin**  
(Antifungal agent)

**Fig. 27** Bioactive natural products containing cyclopropane motifs

Most of these natural products have bioactive properties. Given that fluorine is often a relevant moiety for the control of the chemical properties in drugs, the St Andrews group introduced  $\alpha,\beta,\beta$ -trifluorocyclopropanes as a new fluorinated motif in 2018 to act as a substituent for consideration in drug discovery.<sup>64</sup>

The general synthesis of  $\alpha,\beta,\beta$ -trifluorocyclopropanes involves fluorobromination of a styrene derivative with NBS and HF, followed by a dehydrobromination reaction to yield the corresponding  $\alpha$ -fluorostyrene. These intermediates are treated with  $\text{TMSCF}_3/\text{NaI}$ , which generates difluorocarbene, to afford the final products, as shown in **Scheme 26**.<sup>64</sup>



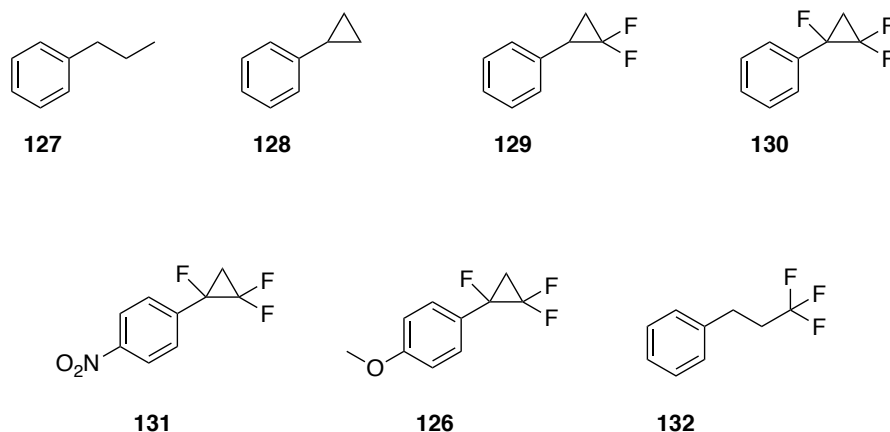
**Scheme 26.** Synthetic procedure to  $\alpha,\beta,\beta$ -trifluorocyclopropanes as reported by C. Thomson<sup>64</sup>

**Scheme 26** presents the syntheses of **125** and **126** as carried out by C. Thomson (University of St Andrews).<sup>64</sup> Samples of these were kindly donated to investigate their biochemical degradation and properties.

### 2.3.2 Aims and Objectives

Cyclopropane moieties are popular motifs in drug discovery programmes. The aim of this section is to study the preliminary viability of the new fluorocyclopropyl moiety in drug discovery programmes. For this, the metabolic fate of **125** and **126** as well as their lipophilicity profiles was studied.

Additionally, selected *n*-propyl and cyclopropyl benzene derivatives were used as tools for a better understanding of the lipophilicity behaviour in these motifs (**Fig. 28**).



**Fig. 28** Selected structures for the study of lipophilicities

### 2.3.3 Estimation of lipophilicities by the use of a reverse-phase HPLC

Log *P* values for the cyclopropyl derivatives **125** and **126** were calculated using the same reference compound set and conditions, similar to that for the thio- and oxy-ether derivatives in **Section 2.2.3**.

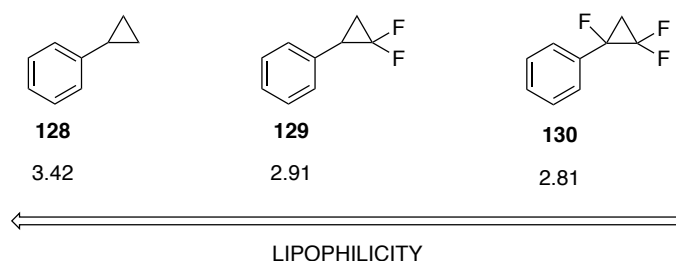
In those cases, the regression line equation was  $y = 0.3772x - 0.493$  ( $R^2 = 0.98572$ ). Using this equation, new Log *P* values were estimated. The experimental results are shown in **Table 8**.

**Table 8.** Retention times measured in triplicates, that led to the estimation of the Log *P* values in blue

Compound	Log <i>P</i>	Rt 1	Rt 2	Rt 3	Average Rt	Capacity factor (k)	Logk
<b>127</b>	<b>4.01</b>	22.45	22.42	22.47	22.45 ± 0.03	10.42	<b>1.02</b>
<b>128</b>	<b>3.42</b>	14.29	14.28	14.31	14.29 ± 0.02	6.27	<b>0.80</b>
<b>129</b>	<b>2.91</b>	9.86	9.86	9.89	9.87 ± 0.02	4.02	<b>0.61</b>
<b>130</b>	<b>2.81</b>	9.21	9.21	9.23	9.22 ± 0.01	3.69	<b>0.57</b>
<b>131</b>	<b>2.79</b>	9.09	9.09	9.11	9.10 ± 0.01	3.63	<b>0.56</b>
<b>126</b>	<b>2.87</b>	9.62	9.63	9.63	9.63 ± 0.01	3.90	<b>0.59</b>
<b>132</b>	<b>3.55</b>	15.76	15.77	15.78	15.77 ± 0.01	7.03	<b>0.85</b>

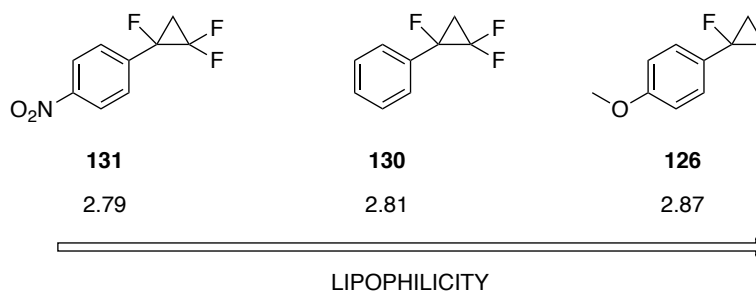
In an attempt to facilitate a comparison between the different compounds and their lipophilicity values, they have been organised into three series. The first series analyses the influence of the different degrees of fluorination upon cyclopropylbenzene **141**, and the second focuses on the effect of an electron-withdrawing and an electron-donating substituents on the aromatic ring. The last series compares the lipophilicity values of alkyl chains and their cyclic analogues, with and without fluorine substitution.

Cyclopropylbenzene **128** had a measured  $\log P$  of 3.42, while the  $\alpha,\beta,\beta$ -trifluorocyclopropyl derivative **130**'s was 2.81. The difluoro cyclopropane **129** falls in-between these at 2.91. However, the value of the latter (**129**) is much closer to the trifluorocyclopropane **130** than to the original cyclopropane **128**, thus the benzylic fluorine does not seem to have a very large effect on the lipophilicity, although its introduction slightly increases rather than decreases polarity.



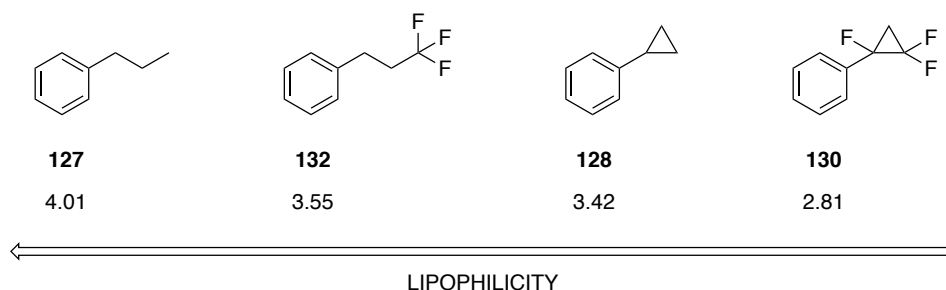
**Fig. 29** Series 1,  $\log P$  comparisons of different degrees of fluorination on cyclopropylbenzene

Given the different electronic properties of electron-withdrawing (EWD) and electron-donating (ED) substituents, **131** and **126** would be expected to have different polarities. However, as shown in **Fig. 30**, it is evident that the difference from the most polar (nitro substituted product **131**) to the least polar (methoxy substituted **126**) is small ( $\Delta \log P = 0.08$ ). This outcome contrasts with the considerable difference in the reported values for nitrobenzene ( $\log P = 1.86$ )<sup>65</sup> and anisole ( $\log P = 2.11$ )<sup>66</sup>; indicating that the effect of the fluorine substitutions on the cyclopropane is larger than that of the aromatic substituents.



**Fig. 30** Series 2, log *P* influence of different substituents on  $\alpha,\beta,\beta$ -trifluorocyclopropyl benzene

In Series 3 (**Fig. 31**), the *n*-propyl substituent has the highest lipophilicity, compared to that of the cyclopropyl derivatives, in both the non-fluorinated and the fluorinated analogues. In each case, partial fluorination lowers the lipophilicity, indicating that fluorination is polarising the vicinal hydrogens. Although trifluorination of *n*-propylbenzene lowers the lipophilicity considerably, cyclopropyl benzene **128** is significantly more polar again. Hence, these partially fluorinated cyclopropyl derivatives become interesting for drug discovery programmes, especially when taking into account that their lipophilic range is below the much desired  $\text{Log } P < 3$  for the trifluorinated analogue **130**.

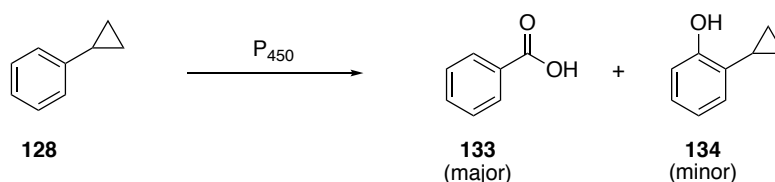


**Fig. 31** Series 3, comparison of cyclic and open chains of propanes, and their partially fluorinated derivatives

### 2.3.4 Metabolic fate of partially fluorinated cyclopropanes in *Cunninghamella elegans*

Cyclopropanes are known to metabolise when exposed to cytochrome P<sub>450</sub> enzymes. Oxidation of cyclopropyl benzene **128** by cytochrome P<sub>450</sub> enzymes was reported by Suckling *et al.* in 1982.<sup>67</sup> In that work, carried out with a rat liver microsomal preparation, 2-cyclopropylphenol **134** was generated. However, this metabolite only

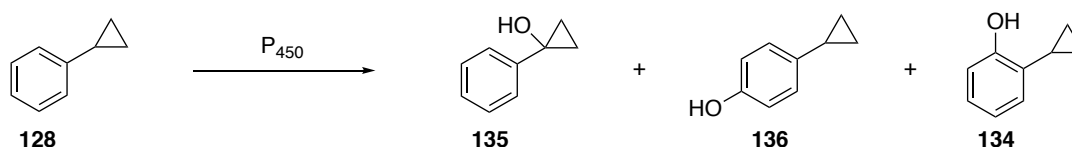
accounts for up to 10% of the products recovered. The major product was benzoic acid **133**.<sup>67,68</sup>



**Scheme 27.** Biotransformations of cyclopropylbenzene **131** by cytochrome P<sub>450</sub> enzymes<sup>67,68</sup>

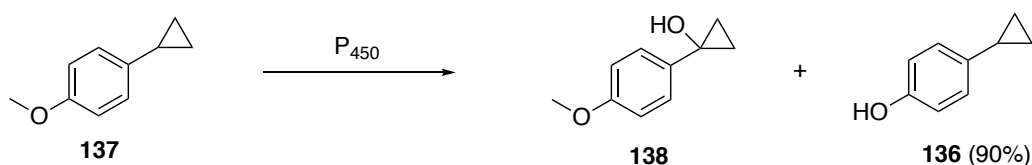
The isolation of benzoic acid **133** indicates the potential for a high level of metabolic activity, and suggested the involvement of a series of oxidative enzymes. Model chemical oxidations carried out on cyclopropylbenzene with cumene hydroperoxide also afforded benzoic acid as a significant product.<sup>68</sup>

Riley and Hanzlik have explored the oxidative metabolism of cyclopropylbenzene **128** in three different microsomal P<sub>450</sub> systems.<sup>69</sup> All three systems gave 1-phenylcyclopropanol **135** by benzylic oxidation, as well as cyclopropyl-4-phenol **136**.<sup>69</sup> One of the systems afforded cyclopropyl-2-phenol **134**, as previously described by Suckling, but none of the microsomes led to the observation of any ring-opened metabolites.<sup>69</sup>



**Scheme 28.** Reported microsomal metabolites of cyclopropylbenzene **131**<sup>69</sup>

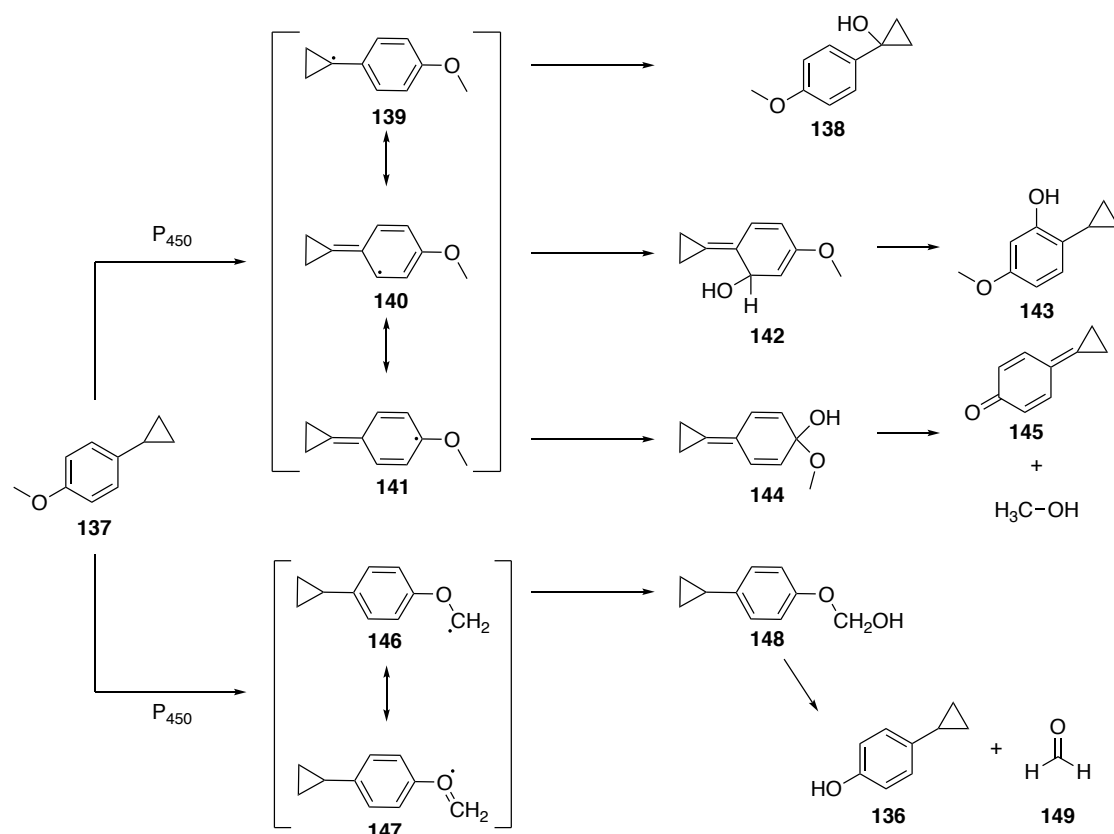
The same three microsomal systems were used to incubate *p*-cyclopropylanisole (**137**), and both an *O*-demethylation product **136** and the benzylic oxidation metabolite **138** were observed. Again, there was no sign of ring opening.<sup>69</sup>



**Scheme 29.** Cytochrome P<sub>450</sub> biodegradation of *p*-cyclopropylanisole **137**<sup>69</sup>

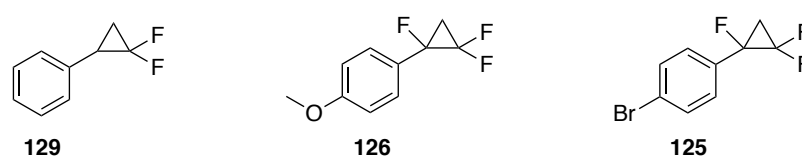
The analogous chemical experiment, carried out by anodic oxidation, did not show any products.<sup>70</sup>

Based on these studies, Vermeulen *et al.* rationalised the origin of the metabolites from cytochrome P<sub>450</sub> oxidation.<sup>71</sup> The computationally obtained results are shown in **Scheme 30**.



**Scheme 30.** Rationale of cytochrome P<sub>450</sub> products for *p*-cyclopropylanisole **137**<sup>71</sup>

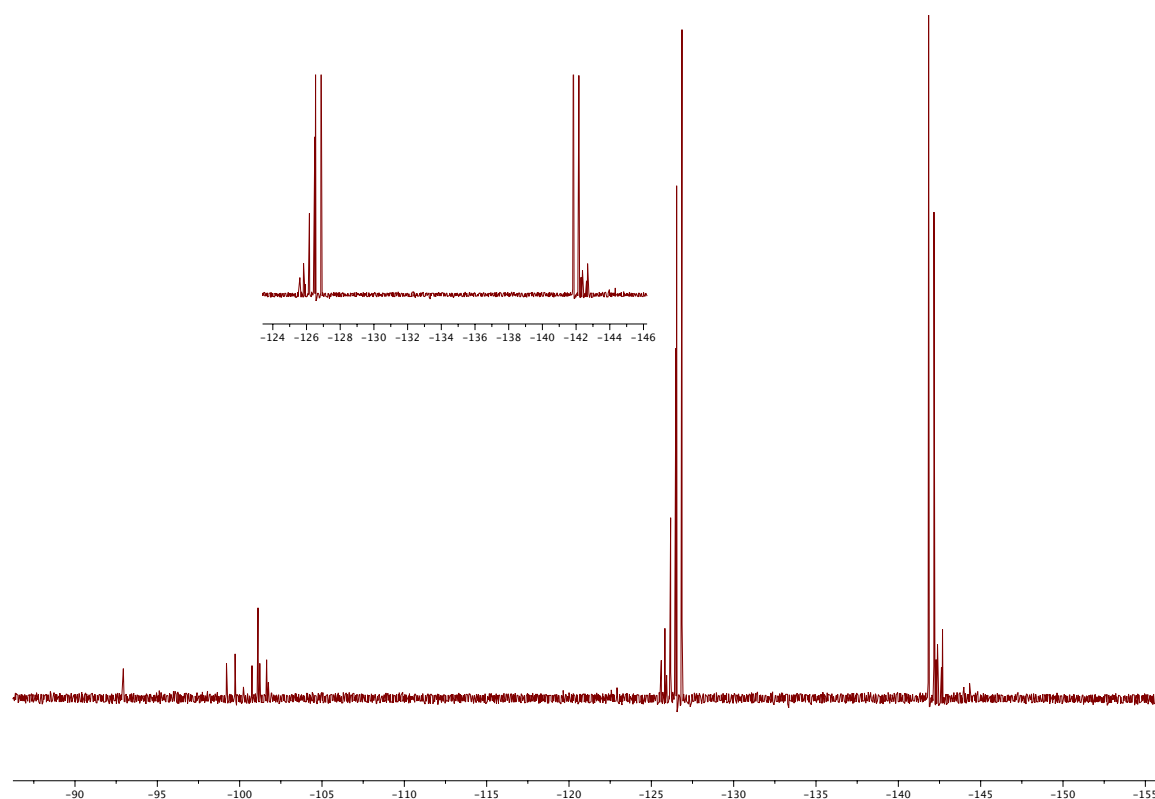
In this Chapter, the metabolism on  $\alpha,\beta,\beta$ -trifluorocyclopropanes was explored to establish the influence of the number of fluorine atoms upon biotransformation, and also to investigate whether different aryl substituents would alter the outcome. The three compounds (**125**, **126** and **129**) shown in **Fig. 32** were selected.



**Fig. 32** Selected fluorocyclopropanes for metabolism studies

#### 2.3.4.1 (2,2-Difluorocyclopropyl)benzene (**129**)

The incubation of cyclopropane **129** with *C. elegans* was carried out over 72 h, at 28 °C and 150 rpm, as previously described. Analysis of  $^{19}\text{F}\{^1\text{H}\}$  NMR spectrum of the cell extract showed that the starting material predominated. However, there were some minor fluorine peaks at a similar chemical shift and with the same coupling pattern, indicating modification of the aromatic ring only. Another product with a very different chemical shift (at around 100 ppm) indicated modification of the  $\beta$ ,  $\beta$ -difluorocyclopropyl motif (**Fig. 33**).



**Fig. 33**  $^{19}\text{F}\{^1\text{H}\}$  NMR spectrum after incubation of **129** with *C. elegans*

Given the very low level of conversion from 10 mg of starting material, no products could be successfully isolated by reverse-phase HPLC.

In order to test the volatility of the metabolites, and whether that would account for the poor recovery, a sample of cyclopropane **129** was left to incubate at 28 °C. Analysis of the crude extracts showed very low intensity signals by fluorine NMR spectroscopy.



This indicated that the fluorometabolites are quite volatile and isolation from the liquid culture is challenging.

A scale-up was investigated. In this case, four different *C. elegans* cultures with 25 mg each of **129** were set up for incubation. However,  $^{19}\text{F}\{^1\text{H}\}$  NMR spectroscopy of the combined crude extracts did not show any traces of fluorine. In the end, high volatility resulted in an incomplete study in this case.

#### **2.3.4.2 1-Bromo-4-(1,2,2-trifluorocyclopropyl)benzene (125)**

1-Bromo-4-(1,2,2-trifluorocyclopropyl)benzene (**125**) was investigated next. After incubation with *C. elegans*, an analysis of the crude extract showed the formation of a minor metabolite. It was clear from the  $^{19}\text{F}$  NMR spectrum that the  $\alpha,\beta,\beta$ -trifluoro motif had remained intact.

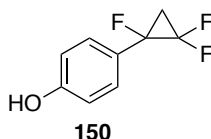
Replicates of this assay showed consistent formation of that same minor metabolite, but attempts of isolation by HPLC proved unsuccessful.

Purification was then explored by column chromatography. In this way, the starting material **125** and the minor metabolite were recovered. However, the latter could only be observed by  $^{19}\text{F}$  NMR spectroscopy, given the low levels that were produced.

Replications did not improve the recovery and the metabolite was not identified.

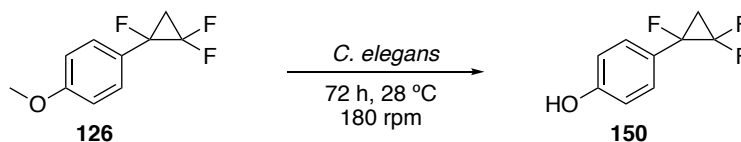
#### **2.3.4.3 1-Methoxy-4-(1,2,2-trifluorocyclopropyl)benzene (126)**

Cultures of *C. elegans* were inoculated with 1-methoxy-4-(1,2,2-trifluorocyclopropyl)benzene (**126**, 5-10 mg). Analysis of the resultant extracts after three days by  $^{19}\text{F}\{^1\text{H}\}$  NMR spectroscopy showed the formation of a new metabolite, along with residual starting material. The  $^{19}\text{F}\{^1\text{H}\}$  NMR patterns revealed that the trifluorocyclopropyl motif was intact and that the modification was to the aromatic ring. Purification by column chromatography allowed product **150** to be isolated.



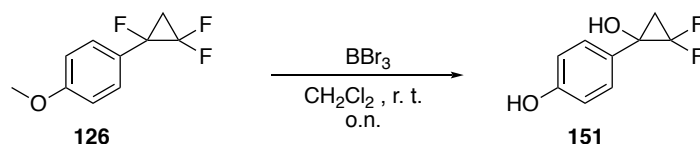
**Fig. 34** 1-methoxy-4-(1,2,2-trifluorocyclopropyl)benzene (**150**)

The biotransformation with *C. elegans* was repeated, increasing the amount of cyclopropane **126** (20 mg), to facilitate the isolation and characterisation of the unknown fluorometabolite. Full NMR analysis, and high resolution mass spectrometry, confirmed the identity of the metabolite as 4-(1,2,2-trifluorocyclopropyl)phenol (**150**).



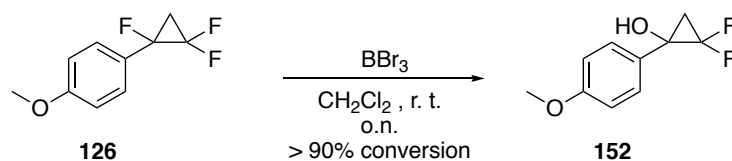
**Scheme 31.** Biotransformation carried out by *C. elegans*

However, previous attempts at the synthesis of **150** had proven unsuccessful (Connor Thompson, University of St Andrews). In that case, the experiment was carried out by treatment of 1-methoxy-4-(1,2,2-trifluorocyclopropyl)benzene (**126**) with boron tribromide (**Scheme 32**). In this experiment, only phenol **151** was isolated. This result seemed to indicate that phenol **150** is unstable, certainly in the presence of  $\text{BBr}_3$ .



**Scheme 32.** Conversion of **126** to **151** by  $\text{BBr}_3$

In order to try to obtain a sample of **150** to secure product identification, it seemed appropriate to repeat this experiment. In my hands, a different result was observed, as illustrated in **Scheme 33**. This generated the cyclopropanol **152**, avoiding the methoxy ether cleavage. Clearly, the  $\text{BBr}_3$  promotes C-F bond cleavage. However, a synthesis of **150** proved elusive.



**Scheme 33.** Reproduction of previous experiment and new product isolated

Two things became clear. Firstly, the fungus does not metabolise the cyclopropane ring, and secondly, *C. elegans* seems to offer a mild method for demethylating **126**, as previously described by Riley and Hanzlik,<sup>69</sup> and Vermeulen.<sup>71</sup>

### 2.3.5 Conclusions on $\alpha,\beta,\beta$ -(trifluorocyclopropyl) aryl derivatives

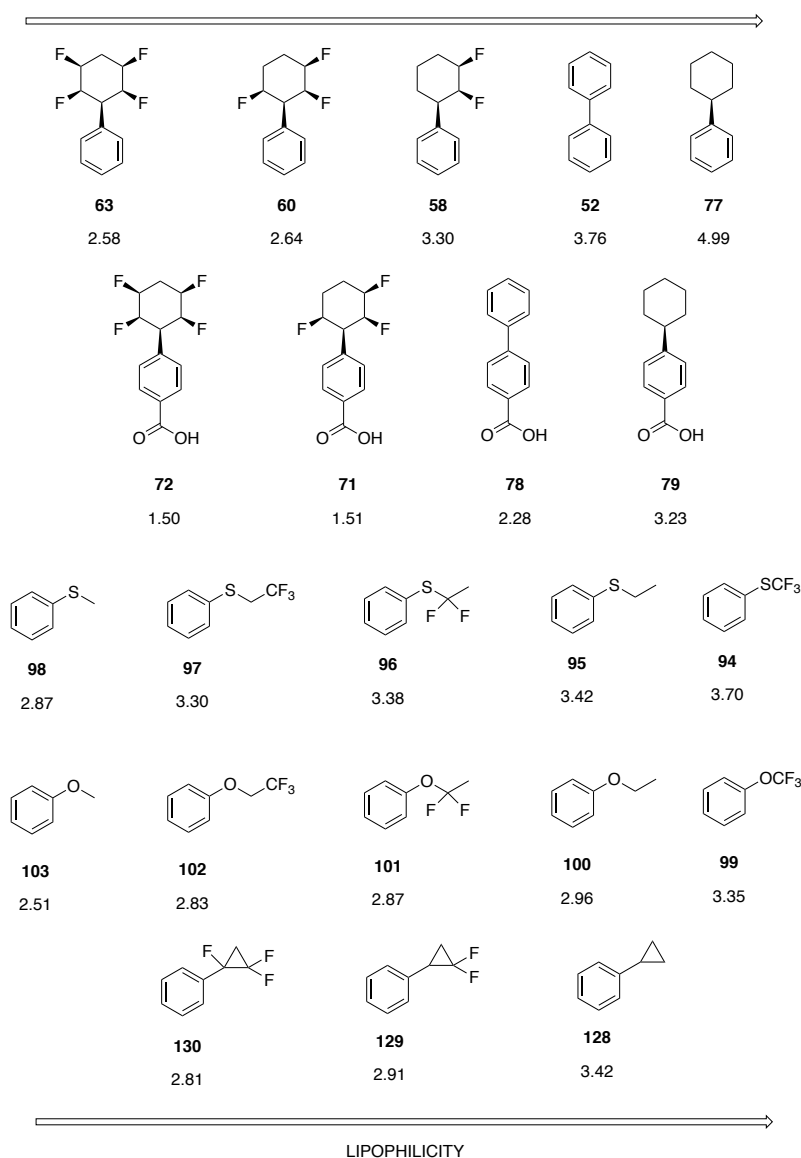
Lipophilicity studies with aryl  $\alpha,\beta,\beta$ -(trifluorocyclopropyl) derivatives led to several observations. Firstly, open chain *n*-propylbenzene is more lipophilic than (cyclopropyl)benzene, and the same tendency is observed for their  $\alpha,\beta,\beta$ -trifluoro analogues. This decrease in Log *P* makes the cyclopropyl moiety attractive as a candidate motif for drug discovery. Additionally, partial substitutions of fluorines for hydrogens increases the polarity further, lowering the lipophilicity values to the 2.50 – 3.00 range.

Metabolism studies with difluorinated aryl (cyclopropyl) derivatives were challenging due to the volatility of the compounds. Most successful was the incubation of 1-methoxy-4-(1,2,2-trifluorocyclopropyl) benzene **126** with *C. elegans*, which led to the formation of 4-(1,2,2-trifluorocyclopropyl)phenol **150**. Notably, the cyclopropyl ring was not metabolised.

## 2.4 General Conclusions on Chapter 2

The main objectives of this Chapter were to investigate the lipophilicity profile and the metabolic stability of a new range of fluorinated motifs recently developed in the St Andrews lab. In total, three new fluorinated motifs were explored: all-*cis* (fluorocyclohexyl) benzenes (**Section 2.1**);  $\alpha,\alpha$ -difluoroethyl thio- and oxy- ethers (**Section 2.2**) and  $\alpha,\beta,\beta$ -(trifluorocyclopropyl) benzenes (**Section 2.3**).

For all of these motifs, the lipophilicity decreased upon the increasing number of fluorine atoms present (**Fig. 35**), making them interesting motifs for drug discovery programmes.



**Fig. 35** Comparison of the Log *P* values obtained for the motifs studied in this Chapter

Metabolic studies with *C. elegans* demonstrated that, most of the products were oxidised, in typical examples of phase I metabolism. In the case of the fluorocyclohexyl derivatives, benzylic hydroxylation was observed in all cases. An increase in fluorination led to a decrease in metabolism of the cyclohexyl ring. The difluorocyclohexane **58** showed a variety of hydroxylation patterns, proving that fluorine substitutions increase metabolic stability.

The  $\alpha,\alpha$ -difluoroethyl thioethers were oxidised on sulfur, either to sulfoxides or sulfones; and a metabolic degradation route for the formation of each metabolite was investigated. Incubation of  $\alpha,\alpha$ -difluoroethyl oxyether (**101**) with *C. elegans* led to the isolation of 4-acetoxyphenol. These results indicate that these oxygen ether motifs are relatively easily metabolised (hydrolysed), and are probably not good candidates for drug discovery programmes, given the very low efficacy they would present with such rapid decomposition.

Finally, the metabolic stability of fluorocyclopropyl derivatives was studied, with different levels of success. While the investigations into (2,2-difluorocyclopropyl)benzene (**129**) and 1-bromo-4-(1,2,2-trifluorocyclopropyl)benzene (**125**) did not lead to definite conclusions; 1-methoxy-4-(1,2,2-trifluorocyclopropyl)benzene (**126**) led to the isolation of demethylated phenol **150**, in a transformation that proved challenging by chemical methods.

Additionally, chiral analysis of some of the products and residual starting material was performed, with varied results. For the cyclohexyl derivatives the enantiomeric excess observed was low (between the 3 – 25% e.e. range); the sulfoxides from the  $\alpha,\alpha$ -difluoroethyl thioethers showed e.e.'s higher than 50%.

## 2.5 References

- 
- <sup>1</sup> S. Purser, P. R. Moore, S. Swallow, V. Gouverneur; *Chem. Soc. Rev.* **2008**, 37, 320.
- <sup>2</sup> H.-J. Böhm, D. Banner, S. Bendels, M. Kansy, B. Kuhn, K. Müller, U. Obst-Sander, M. Stahl; *ChemBioChem*, **2004**, 5, 637.
- <sup>3</sup> E. P. Gillis, K. J. Eastman, M. D. Hill, D. J. Donnelly, N. A. Meanwell; *J. Med. Chem.* **2015**, 58, 8315.
- <sup>4</sup> C. Vraha, L. Nics, K.-H. Wagner, M. Hacker, W. Wadsak, M. Mitterhauser; *NucMedBio*, **2017**, 50, 1.
- <sup>5</sup> B.-C. Wang, L.-J. Wang, B. Jiang, S.-Y. Wang, N. Wu, L.-Q. Li, D.-Y. Shi; *Mini Reviews in Medicinal Chemistry*, **2017**, 17, 683.
- <sup>6</sup> A. Tressaud, G. Haufe; *Fluorine and Health. Molecular Imaging, Biomedical Materials and Pharmaceuticals*, Elsevier (**2008**).
- <sup>7</sup> Q. A. Huchet, B. Kuhn, B. Wagner, H. Fischer, M. Kansy, D. Zimmerli, E. M. Carreira, K. Müller; *J. Fluorine Chem.* **2013**, 152, 119.
- <sup>8</sup> W. Palmer-Brown, B. Dunne, Y. Ortin, M. A. Fox, G. Sandford, C. D. Murphy; *Xenobiotica*, **2017**, 47, 763.
- <sup>9</sup> Y. W. J. Wong, P. J. Davis; *Pharm. Res.* **1989**, 6, 982.
- <sup>10</sup> G. P. Rao, P. J. Davis; *Drug Metab. Dispos.* **1997**, 25, 709.
- <sup>11</sup> D. Zhang, Y. Yang, J. E. A. Leakey, C. E. Cerniglia; *FEMS Microbiol. Lett.* **1996**, 138, 221.
- <sup>12</sup> D. H. R. Barton; *Experientia*, **1950**, 8, 316.
- <sup>13</sup> a) O. Hassel, B. Ottar; *Arch. Math. Naturvidenskab*, **1942**, XLV. b) O. Hassel, H. Viervoll; *Acta Chemica Scandinavica*, **1947**, 1, 149. c) O. Hassel, B. Ottar; *Acta Chemica Scandinavica*, **1947**, 1, 929.
- <sup>14</sup> The Nobel Prize in Chemistry 1969. Nobelprize.org. Nobel Media AB 2018.

- 
- <sup>15</sup> A. Schatz, E. Bugie, S. A. Waksman; *Proc. Soc. Exper. Biol. Med.* **1944**, *55*, 66.
- <sup>16</sup> S. A. Waksman, H. A. Lechevalier; *Science*, **1949**, *109*, 305.
- <sup>17</sup> L. Wanka, K. Iqbal, P. R. Schreiner; *Chem. Rev.* **2013**, *113*, 3516.
- <sup>18</sup> M. Aldeghi, S. Malhotra, D. L. Selwood, A. W. E. Chan; *Chem. Biol. Drug Des.* **2014**, *83*, 450.
- <sup>19</sup> A. J. Durie, A. M. Z. Slawin, T. Lebl, P. Kirsch, D. O'Hagan; *Chem. Commun.* **2011**, *47*, 8265.
- <sup>20</sup> A. J. Durie, A. M. Z. Slawin, T. Lebl, P. Kirsch, D. O'Hagan; *Chem. Commun.* **2012**, *48*, 9643.
- <sup>21</sup> N. S. Keddie, A. M. Z. Slawin, T. Lebl, D. Philp, D. O'Hagan; *Nature Chemistry*, **2015**, *7*, 483.
- <sup>22</sup> M. P. Wiesenfeldt, Z. Nairoukh, W. Li, F. Glorius, *Science*, **2017**, *357*, 908.
- <sup>23</sup> M. P. Wiesenfeldt, T. Knecht, C. Schlepphorst, F. Glorius; *Angew. Chem. Int. Ed.* **2018**, *57*, 8297.
- <sup>24</sup> A. J. Durie, T. Fujiwara, R. Cormanich, M. Bühl, A. M. Z. Slawin, D. O'Hagan; *Chem. Eur. J.* **2014**, *20*, 6259.
- <sup>25</sup> A. J. Durie, T. Fujiwara, N. Al-Maharik, A. M. Z. Slawin, D. O'Hagan; *J. Org. Chem.* **2014**, *79*, 8228.
- <sup>26</sup> A. Rodil, S. Bosisio, M. S. Ayoup, L. Quinn, D. B. Cordes, A. M. Z. Slawin, C. D. Murphy, J. Michel, D. O'Hagan; *Chem. Sci.* **2018**, *9*, 3023.
- <sup>27</sup> M. S. Ayoup, D. B. Cordes, A. M. Z. Slawin, D. O'Hagan; *Beilstein J. Org. Chem.* **2015**, *11*, 2671.
- <sup>28</sup> C. A. Lipinski, F. Lombardo, B. W. Dominy, P. J. Feeney; *Adv. Drug Deliv. Rev.* **1997**, *23*, 3.
- <sup>29</sup> C. A. Lipinski; *Adv. Drug Deliv. Rev.* **2016**, *101*, 34.

- 
- <sup>30</sup> a) J. Sangster, *J. Phys. Chem. Ref. Data*, **1989**, *18*, 1111. b) C. Hansch, A. Leo, D. Hoekman, *Exploring QSAR – Hydrophobic, Electronic, and Steric Constants*, Washington, DC: American Chemical Society, **1995**. c) P. C. von der Ohe, R. Kühne, R.-U. Ebert, R. Altenburger, M. Liess, G. Schüürmann, *Chem. Res. Toxicol.* **2005**, *18*, 536 – 555.
- <sup>31</sup> B. Linclau, Z. Wang, G. Compain, V. Paumelle, C. Q. Fontenelle, N. Wells, A. Weymouth-Wilson; *Angew. Chem. Int. Ed.* **2016**, *55*, 674.
- <sup>32</sup> C. Gianinis, A. Tsantili-Kakoulidou, *J. Liq. Chromatogr. Relat. Technol.* **2008**, *31*, 79.
- <sup>33</sup> S. Bosisio, A. S. J. S. Mey, J. Michel; *J. Comput. Aided Mol. Des.* **2016**, *30*, 1101.
- <sup>34</sup> T. V. Bright, F. Dalton, V. L. Elder, C. D. Murphy, N. K. O'Connor, G. Sandford; *Org. Biomol. Chem.* **2013**, *11*, 1135.
- <sup>35</sup> M. J. Shaughnessy, A. Harsanyi, J. Li, T. V. Bright, C. D. Murphy, G. Sandford; *ChemMedChem*, **2014**, *9*, 733.
- <sup>36</sup> T. Bykova, N. Al-Maharik, A. M. Z. Slawin, M. Bühl, T. Lebl, D. O'Hagan; *Chem. Eur. J.* **2018**, *24*, 13290.
- <sup>37</sup> P. Dawar, M. B. Raju, R. A. Ramakrishna; *Synth. Commun.* **2014**, *44*, 836.
- <sup>38</sup> R. Tomita, N. Al-Maharik, A. Rodil, M. Bühl, D. O'Hagan; *Org. Biomol. Chem.* **2018**, *16*, 1113.
- <sup>39</sup> F. Toulgoat, S. Alazet, T. Billard; *Eur. J. Org. Chem.* **2014**, 2415.
- <sup>40</sup> H. Jiang, R. Zhu, C. Zhu, F. Chen, W. Wu; *Org. Biomol. Chem.* **2018**, *16*, 1646.
- <sup>41</sup> A. Riesco-Domínguez, J. van de Wiel, T. A. Hamlin, B. van Beek, S. D. Lindell, D. Blanco-Ania, F. M. Bickelhaupt, F. P. J. T. Rutjes; *J. Org. Chem.* **2018**, *83*, 1779.
- <sup>42</sup> S. Z. Zard; *Org. Biomol. Chem.* **2016**, *14*, 6891.
- <sup>43</sup> K. Uneyama; *Jpn. Kokai Tokkyo Koho*, **1990**, JP 02121961 A.



- 
- <sup>44</sup> A. Koehler, S. Cerezo-Galvez, B. Alig, R. Fischer, J. J. Hahn, K. Ilg, D. Portz, O. Malsam, P. Loesel, D. Wilcke; *PCT Int. Appl.* **2015**, WO 2015091267 A1.
- <sup>45</sup> P. Jimonet, F. Audiau, M. Barreau, J.-C. Blanchard, A. Boireau, Y. Bour, M.-A. Coléno, A. Doble, G. Doerflinger, C. D. Huu, M.-H. Donat, J. M. Duchesne, P. Ganil, C. Guérémy, E. Honoré, B. Just, R. Kerphirique, S. Gontier, P. Hubert, P. M. Laduron, J. Le Blevec, M. Meunier, J.-M. Miquet, C. Nemecek, M. Pasquet, O. Piot, J. Pratt, J. Rataud, M. Reibaud, J.-M. Stutzmann, S. Mignami; *J. Med. Chem.* **1999**, *42*, 2828.
- <sup>46</sup> N. Coleman, H. M. Nguyen, Z. Cao, B. M. Brown, D. P. Jenkins, D. Zolkowska, Y.-J. Chen, B. S. Tanaka, A. L. Goldin, M. A. Rogawski, I. N. Pessah, H. Wulff; *Neurotherapeutics*, **2015**, *12*, 234.
- <sup>47</sup> a) Q. A. Huchet, B. Kuhn, B. Wagner, N. A. Kratochwil, H. Fischer, M. Kansy, D. Zimmerli, E. M. Carreira, K. Müller; *J. Med. Chem.* **2015**, *58*, 9041. b) K. Müller; *Chimia*, **2014**, *68*, 356.
- <sup>48</sup> Y. Zafrani, D. Yeffet, G. Sod-Moriah, A. Berliner, D. Amir, D. Marciano, E. Gershonov, S. Saphier; *J. Med. Chem.* **2017**, *60*, 797.
- <sup>49</sup> C. L. Coombes; *J. Org. Chem.* **2008**, *73*, 6758.
- <sup>50</sup> W. H. Perkin (junior), *Eur. J. Inorg. Chem.* **1884**, *17*, 54.
- <sup>51</sup> D. Y.-K. Chen, R. H. Pouwer, J.-A. Richard; *Chem. Soc. Rev.* **2012**, *41*, 4631.
- <sup>52</sup> L. A. Wessjohann, W. Brandt; *Chem. Rev.* **2003**, *103*, 1625.
- <sup>53</sup> A. De Meijere, *Chem. Revs.* **2003**, *103* (4). Editorial.
- <sup>54</sup> Y.-Y. Fan, X.-H. Gao, J.-M. Yue; *Sci. China Chem.* **2016**, *59*, 1126.
- <sup>55</sup> Y.-M. Shi, X.-B. Wang, X.-N. Li, X. Luo, Z.-Y. Shen, Y.-P. Wang, W.-L. Xiao, H.-D. Sun; *Org. Lett.* **2013**, *15*, 5068.
- <sup>56</sup> P. C. Wang, X. H. Ran, H. R. Luo, J. M. Chen, Q. Y. Ma, H. F. Dai, Y. Q. Liu, M. J. Xie, J. Zhou, Y. X. Zhao; *Org. Lett.* **2011**, *13*, 3036.
- <sup>57</sup> Y. Tian, Q. Guo, W. Xu, G. Zhu, Y. Yang, J. Shi; *Org. Lett.* **2014**, *16*, 3950.

- 
- <sup>58</sup> Y.-C. Chang, C.-K. Lu, Y.-R. Chiang, G.-J. Wang, Y.-M. Ju, Y.-H. Kuo, T.-H. Lee; *J. Nat. Prod.* **2014**, *77*, 751.
- <sup>59</sup> C.-H. Li, S.-X. Jing, S.-H. Luo, W. Shi, J. Hua, Y. Liu, X.-N. Li, B. Schneider, J. Gershenzon, S.-H. Li; *Org. Lett.* **2013**, *15*, 1694.
- <sup>60</sup> T. Hiratsuka, H. Suzuki, R. Kariya, T. Seo, A. Minami, H. Oikawa; *Angew. Chem. Int. Ed. Engl.* **2014**, *53*, 5423.
- <sup>61</sup> S. Hanessian, U. Soma, S. Dorich, B. Deschênes-Simard; *Org. Lett.* **2011**, *13*, 1048.
- <sup>62</sup> L. Zhu, R. Tong; *J. Antibiot.* **2016**, *69*, 280.
- <sup>63</sup> M. J. Balunas, M. F. Grosso, F. A. Villa, N. Engene, K. L. McPhail, K. Tidgewell, L. M. Pineda, L. Gerwick, C. Spadafora, D. E. Kyle, W. H. Gerwick; *Org. Lett.* **2012**, *14*, 3878.
- <sup>64</sup> C. J. Thomson, Q. Zhang, N. Al-Maharik, M. Bühl, D. B. Cordes, A. M. Z. Slawin, D. O'Hagan; *Chem. Commun.* **2018**, *54*, 8415.
- <sup>65</sup> K. Roy, P. L. A. Popelier; *QSAR Comb. Sci.* **2008**, *8*, 1006.
- <sup>66</sup> H. Sanderson, M. Thomsen; *Toxicol. Lett.* **2009**, *187*, 84.
- <sup>67</sup> K. E. Suckling, C. G. Smellie, I. E. S. Ibrahim, D. C. Nonhebel, C. J. Suckling; *FEBS Lett.* **1982**, *145*, 179.
- <sup>68</sup> C. J. Suckling, D. C. Nonhebel, L. Brown, K. E. Suckling, S. Seilman, C. R. Wolf; *Biochem. J.* **1985**, *232*, 199.
- <sup>69</sup> P. Riley, R. P. Hanzlik; *Xenobiotica*, **1994**, *24*, 1.
- <sup>70</sup> T. Shono, Y. Matsumara; *J. Org. Chem.* **1970**, *35*, 4157.
- <sup>71</sup> M. J. de Groot, G. M. Donné-Op den Kelder, J. N. M. Commandeur, J. H. van Lenthe, N. P. E. Vermeulen; *Chem. Res. Toxicol.* **1995**, *8*, 437.

## Chapter 3. *In vitro* reconstitution of (2*R*,3*S*,4*S*)-5-fluoro-2,3,4-trihydroxypentanoic acid biosynthesis

### 3.1 Introduction to (2*R*,3*S*,4*S*)-5-fluoro-2,3,4-trihydroxypentanoic acid (**2**) biosynthesis

In 2014, the St. Andrews group reported the isolation and characterisation of a new fluorinated natural product: (2*R*,3*S*,4*S*)-5-fluoro-2,3,4-trihydroxypentanoic acid **2** (5-FHPA).<sup>1</sup> This was enabled by the discovery of a fluorinase in *Streptomyces* sp. MA37, after a full genome sequence.<sup>2,3</sup>

The fluorometabolite profile of *S.* sp. MA37 is more complex than that of *Streptomyces cattleya*. Fluoroacetate **1** and 4-fluorothreonine **4** are produced, but there was also a number of unknown fluorinated metabolites, as can be seen in the <sup>19</sup>F NMR spectrum (Fig. 1).<sup>1</sup>

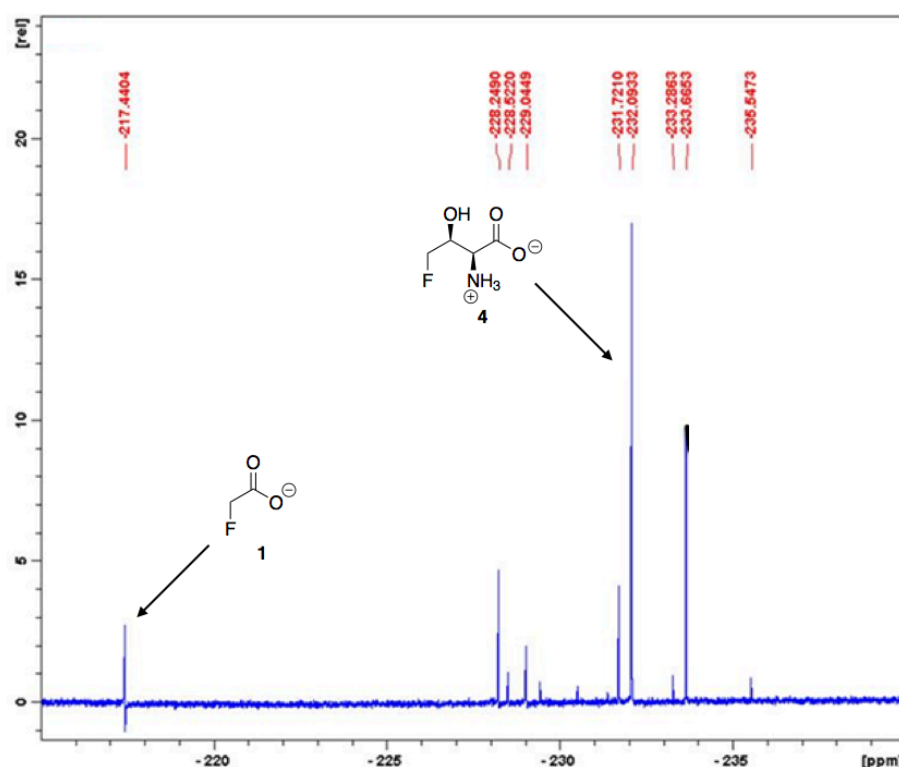
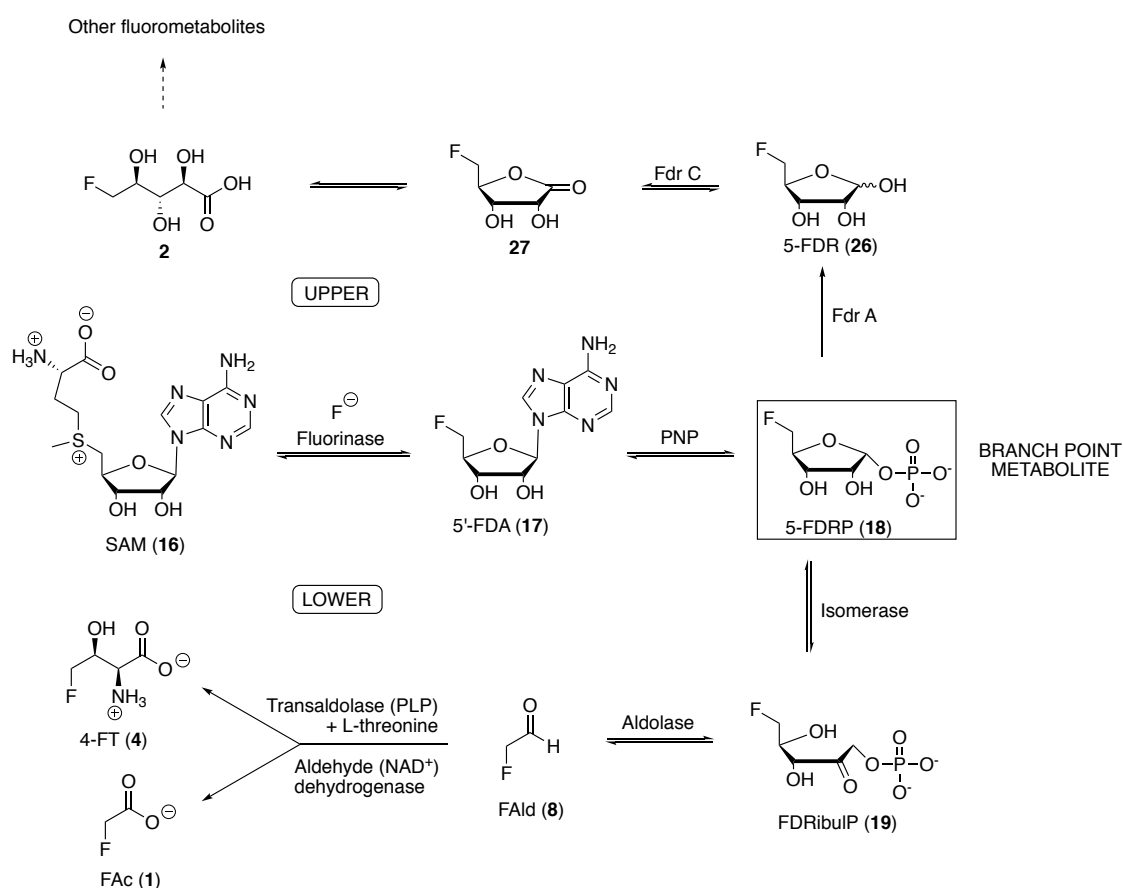


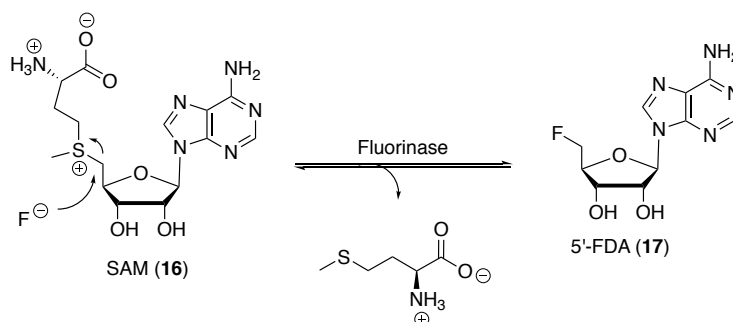
Fig. 1 <sup>19</sup>F{<sup>1</sup>H} NMR spectrum of the fluorometabolite profile of *S.* sp. MA37<sup>1</sup>

The first steps of the metabolic pathway of *S. sp.* MA37 are similar to those of *S. cattleya*.<sup>1</sup> However, the metabolic pathway in *S. sp.* MA37 branches into two. The first branch leads to the formation of fluoroacetate (**1**) and 4-fluorothreonine (**4**); but the second generates 5-FHPA (**2**) and other unknown metabolites. Both branches are shown in **Scheme 1**, indicated as ‘upper’ and ‘lower’.



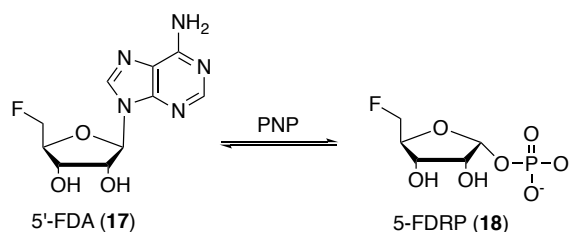
**Scheme 1.** Biosynthetic pathway to the fluorometabolites in *S. sp.* MA37

The process that allows the formation of 5-FHPA **2** requires four enzymes.<sup>1</sup> It starts with a fluorinase reaction on SAM to generate 5'-FDA (**Scheme 2**).<sup>4</sup> The fluorinase in *S. sp.* MA37 (*flA1*) has a gene with an 87% homology to that from *S. cattleya*.<sup>2</sup> For expression in *E. coli*, a synthetic gene for *flA1* was codon-optimised and over-expressed. Additionally, a His<sub>6</sub> tag and a TEV protease cleavage site were added into the synthetic gene, in order to facilitate the isolation and purification of the recombinant enzyme.<sup>2</sup> Incubation with SAM and fluoride demonstrated an active fluorinase.<sup>2</sup>



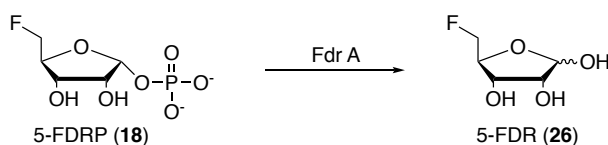
**Scheme 2.** Fluorinase-catalysed formation of 5'-FDA (17)<sup>4</sup>

In the metabolic pathway, 5'-FDA is then transformed into 5-FDRP, in a step catalysed by a purine nucleoside phosphorylase (PNP).<sup>5</sup> PNPs are well known to cleave nucleoside bases and to generate the corresponding ribose phosphates; therefore, phosphate ions must be available in the medium.<sup>6</sup>



**Scheme 3.** Step 2, conversion of 5'-FDA into 5-FDRP by PNP<sup>5</sup>

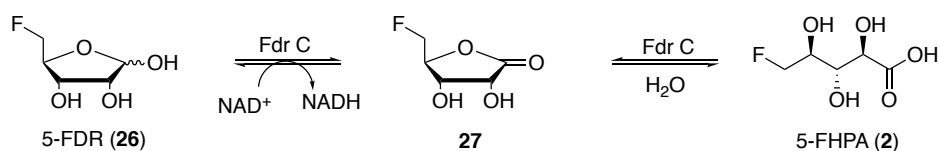
5-FDRP is a branch point of the metabolic pathway. The metabolite can now either follow the pathway reported for *S. cattleya* to generate fluoroacetate and 4-fluorothreonine<sup>7</sup>, or ribose phosphate **18** can be transformed by a dephosphorylase (FdrA) into 5-FDR **26** (**Scheme 4**), in an analogous way to that of chlororibose in *Salinispora tropica*, as discussed in Chapter 1 (**Section 1.5.3**).<sup>1,8</sup> FdrA has been predicted to encode a metal-dependent phosphoesterase, with a 56% identity to that found in *S. tropica*, but the information available on this protein is limited at present.<sup>1,8</sup>



**Scheme 4.** FdrA dephosphorylates 5-FDRP to yield 5-FDR

The final enzyme involved in this pathway to **2** is FdrC, which catalyses an oxidation followed by a hydrolysis. Firstly, it transforms 5-FDR into lactone **27**, which is then

hydrolysed to the open chain carboxylic acid 5-FHPA **2**.<sup>1</sup> FdrC is an NAD<sup>+</sup> dependent short chain dehydrogenase, which shares a 69% homology with the *S. tropica* analogue.<sup>1,8</sup> A codon-optimised synthetic gene for FdrC was over-expressed in *E. coli* and purified, a process also facilitated by the addition of a His<sub>6</sub> tag and a TEV protease cleavage site to the gene.<sup>1</sup>



**Scheme 5.** Final biotransformation into 5-FHPA (**2**) via lactone intermediate **27**<sup>1</sup>

## 3.2 Aims and Objectives

The main objective of this Chapter is to reproduce the *in vitro* biotransformations that lead to the formation of 5-FHPA **2** from SAM and fluoride.

Additionally, as well as reconstituting the biosynthetic pathway, optimisation of the conditions will be explored, in order to attempt a milligram scale-up in the production of 5-FHPA (**2**).

## 3.3 *In vitro* reconstitution of 5-FHPA biosynthesis

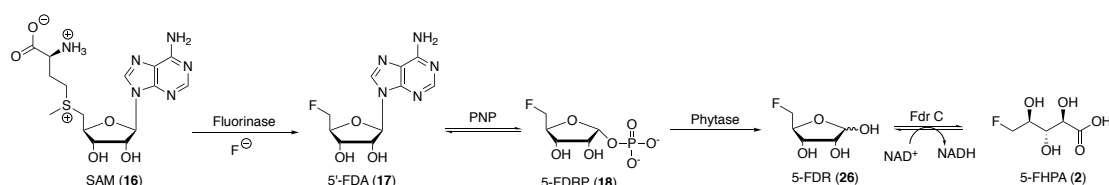
### 3.3.1 Transformation of fluoride to 5-FHPA (**2**)

The *in vitro* reconstitution of the 5-FHPA biosynthesis was inspired by a previous reconstitution of the formation of 4-fluorothreonine from fluoride in a one-pot, five-enzyme experiment in an NMR tube.<sup>9</sup>

In order to reproduce the natural biosynthesis of 5-FHPA *in vitro*, it was necessary to obtain the enzymes and substrates involved in each of the four enzymatic steps. The recombinant fluorinase from *S. cattleya* used in the following sections was over-expressed and purified by Dr Stephen Thompson (University of St Andrews).<sup>4</sup> Recombinant PNP from *E. coli* is available from Sigma Aldrich. A phytase from wheat, also available from Sigma Aldrich, was used in this study, instead of FdrA. The oxido-reductase FdrC was required to be over-expressed and this was achieved as

described in **Section 5.6.4 (Chapter 5)**, following a previous protocol.<sup>1</sup> SAM, KF and  $\text{NAD}^+$  are commercially available and were sourced from Sigma Aldrich.

The four-enzyme protocol is illustrated in **Scheme 6**. Accordingly,  $\text{SAM} \cdot 2\text{HCl}$  was incubated with KF,  $\text{NAD}^+$ , fluorinase, PNP, the commercial phytase from wheat and FdrC. The experiment was carried out in PBS buffer, at 37 °C for 16 h.

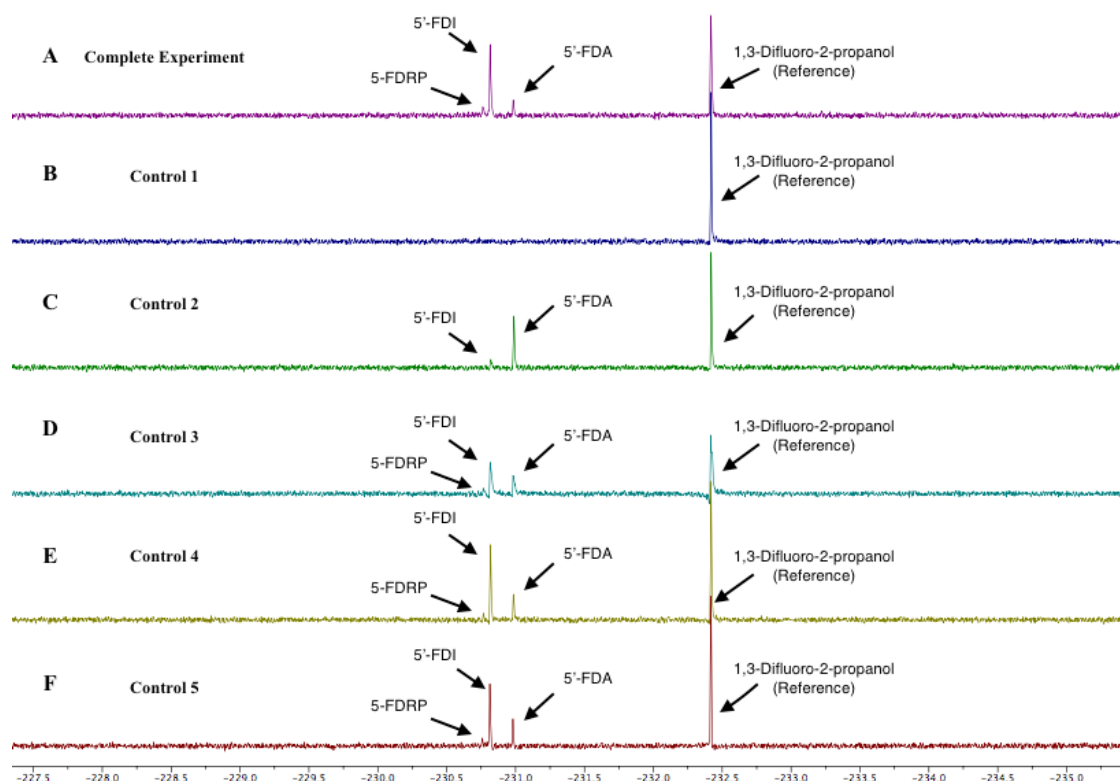


**Scheme 6.** Four-enzyme biotransformation of fluoride to 5-FHPA 2

The  $^{19}\text{F}\{^1\text{H}\}$  NMR spectrum carried out after incubation indicated the formation of three organofluorine products, which were identified as 5'-FDI, 5'-FDA and 5-FDRP, in order of abundance (**Fig. 2A**). The identity of these products was secured by comparison of their  $^{19}\text{F}$  NMR spectra to that from previous published work in the group.<sup>5</sup>

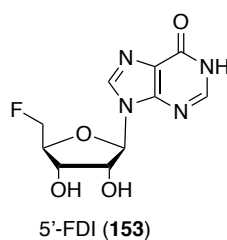
In addition, five control incubations (**Fig. 2B-F**) were carried out to ensure the correct functioning of each enzyme, or to detect if any step is problematic for efficient turnover. In each control, one of the enzymes was removed. Therefore, Control 1 (**Fig. 2B**) was run without fluorinase; Control 2 (**Fig. 2C**) did not contain PNP; Control 3 (**Fig. 2D**) did not have phytase, and Control 4 (**Fig. 2E**) was lacking FdrC. Control 5 (**Fig. 2F**) was carried out without  $\text{NAD}^+$ . The  $^{19}\text{F}$  NMR spectra of these control experiments are shown in **Fig. 2**, in descending order.

It is important to note that most of the fluorine peaks for the products involved in these experiments appear within a 0.5 ppm range from each other. Therefore, it is challenging to assign the identity of the peak without an internal reference. 1,3-Difluoropropanol was chosen for this purpose, given that it is a fluoroorganic compound and its chemical shift is only 2 ppm upfield, but sufficiently remote not to overlap any of the studied products. Additionally, 1,3-difluoropropanol is stable in all the buffers and conditions used.



**Fig. 2** Complete experiment (top) and control experiments 1 to 5 (descending order) for the biotransformation experiments. A) The complete experiment contained all the enzymes and cofactors, as well as fluoride B) Control 1. No fluorinase C) Control 2. No PNP D) Control 3. No phytase E) Control 4. No FdrC F) Control 5. No  $\text{NAD}^+$

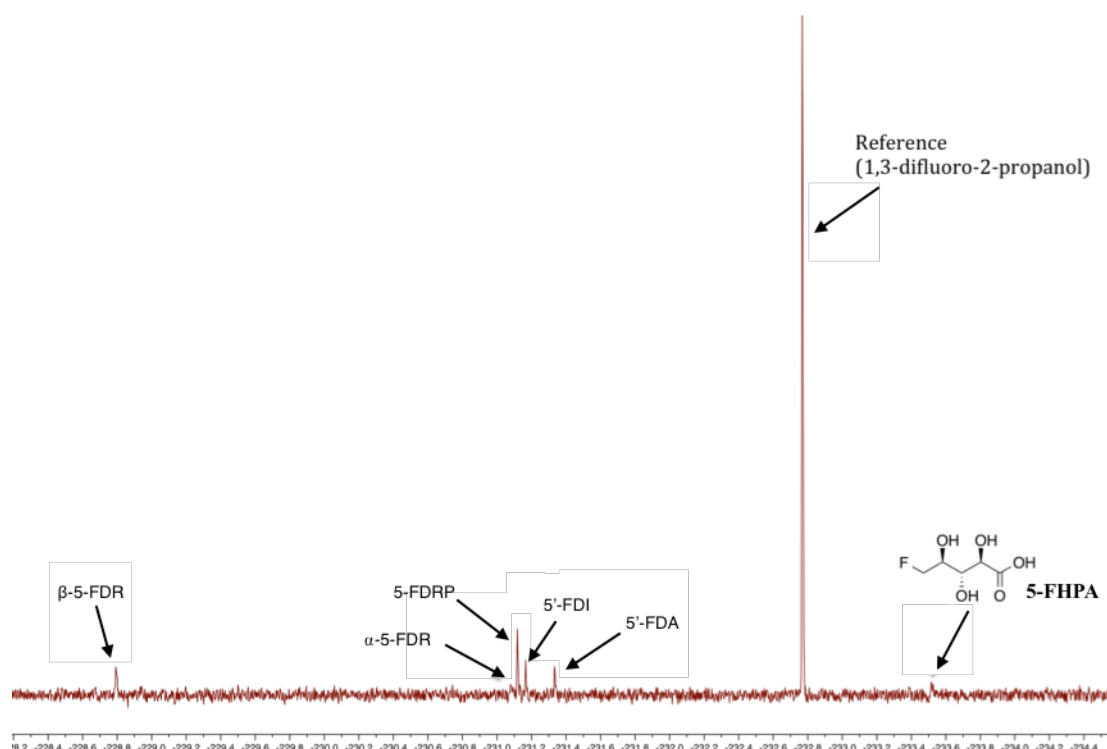
As expected, Control 1 (**B**) showed that without the fluorinase there is no formation of organofluorine compounds. Control 2 (**C**) indicates that the fluorinase is effective in the generation of 5'-FDA, although it shows a minor deaminase activity to yield 5'-FDI (**153**), presumably due to enzymatic contamination.



Controls 3 (**D**), 4 (**E**) and 5 (**F**) lead to the same outcome as that of the 'complete experiment' (**A**) with all the enzymes and cofactors, suggesting that step three (the PNP biotransformation) is blocking the overall biotransformation to 5-FHPA.

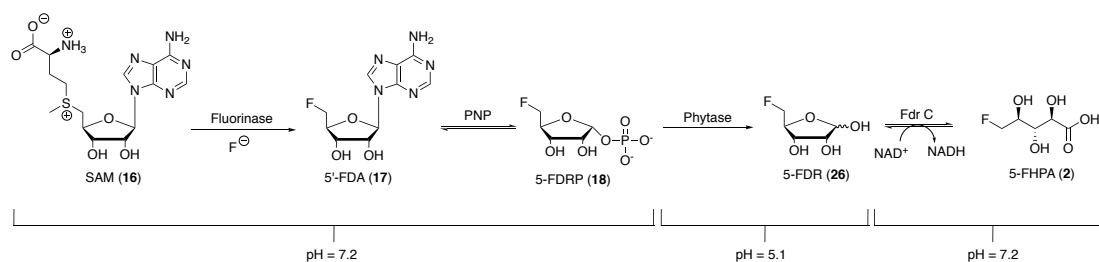


The same recombinant PNP from *E. coli* was sourced from Sigma-Aldrich, to probe whether the stock in the lab had been contaminated or inactive after the time stored, and the complete experiment was repeated. This substitution improved the biotransformation, and in this case did generate 5-FHPA **2**, although **2** is observed at a low level (**Fig. 3**). The major product observed by  $^{19}\text{F}\{^1\text{H}\}$  NMR spectroscopy is 5-FDRP, suggesting that the problematic step is the dephosphorylation reaction catalysed by the phytase.



**Fig. 3**  $^{19}\text{F}\{^1\text{H}\}$  NMR spectrum of the ‘complete experiment’, showing a low formation of 5-FHPA

Although the commercial phytase from wheat seems to have some activity in PBS buffer (pH = 7.4) at 37 °C, the optimal conditions for this enzyme are at pH = 5 and 55 °C.<sup>10</sup> Therefore, a new three-step, one-pot *in vitro* reconstitution was explored. This new assay involved two pH changes, as illustrated in **Scheme 7**.



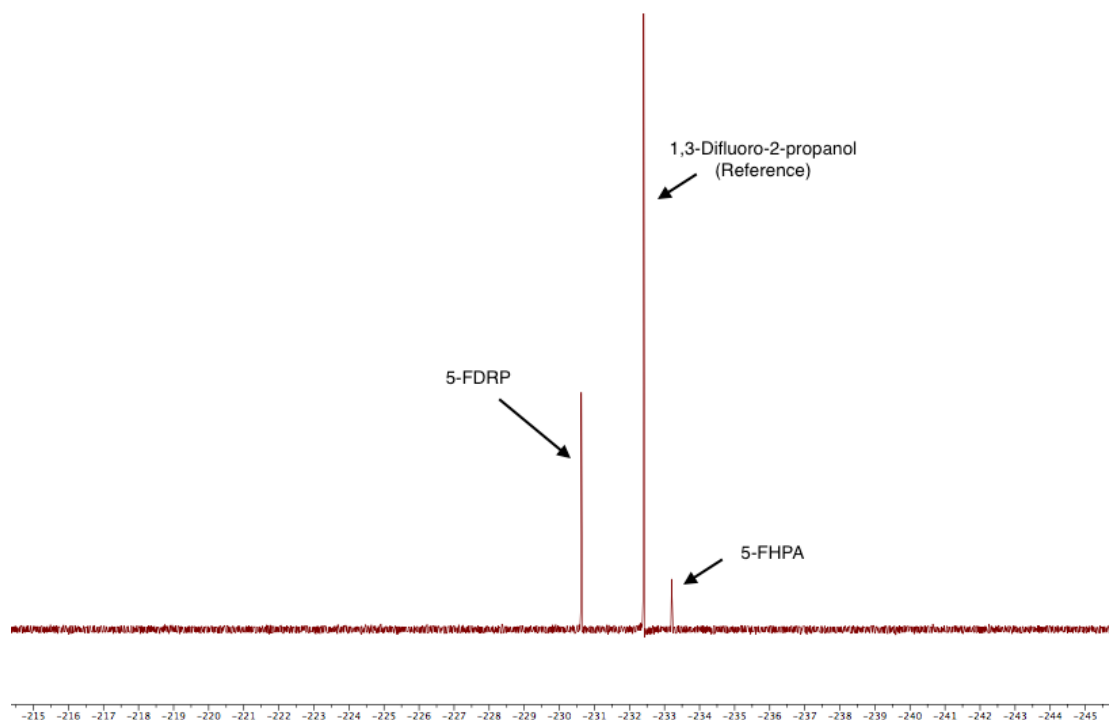
**Scheme 7.** Envisaged 3-steps 1-pot biotransformation to 5-FHPA (2)

Since the aim was to carry out this experiment in just one pot, it became crucial to use a buffer that would allow the necessary changes of pH. A phosphate-citrate buffer was selected. At this stage, the lab received a generous donation of a PNP, immobilised in Sepabeads®, from GlaxoSmithKline (GSK). Unfortunately, the precise origin of this PNP could not be established, although it was a recombinant PNP over-expressed in *E.coli*, and possibly, of human origin. This new PNP proved very successful, and allowed a 100% conversion of 5'-FDA into 5-FDRP (see **Section 3.3.2.2**). Hence, this PNP was used in all future experiments.

The three-step, one-pot experiment described in **Scheme 7** started with the addition of SAM and KF into an eppendorf tube, followed by the buffer (at pH = 7.2). PNP and fluorinase were then added, and the tube was set to incubate at 37 °C for 24 h. A gentle rotation was applied for a better mixture of the assay.

After 24 h, the immobilised PNP was filtered, and the fluorinase was denatured by heating. The precipitated protein was then centrifuged. The clear supernatant was acidified to pH = 5.1 by gentle addition of citric acid. Then, the phytase was added, and the mixture set to incubate at 40 °C for 48 h with gentle rotation.

After the incubation period, the phytase reaction was stopped by raising the pH of the mixture to 7.2, which triggered denaturation of the phytase. The precipitated phytase was filtered by centrifugation, and FdrC and  $NAD^+$  were added to the supernatant. The biotransformation was again incubated at 37 °C for 30 h. A  $^{19}F\{^1H\}$  NMR spectrum of the supernatant was carried out after denaturing the enzyme. The result is shown in **Fig. 4**. This assay was successful in the formation of 5-FHPA. However, the conversion was not complete, and the major product was residual 5-FDRP; indicating that the FdrC step was not efficient.



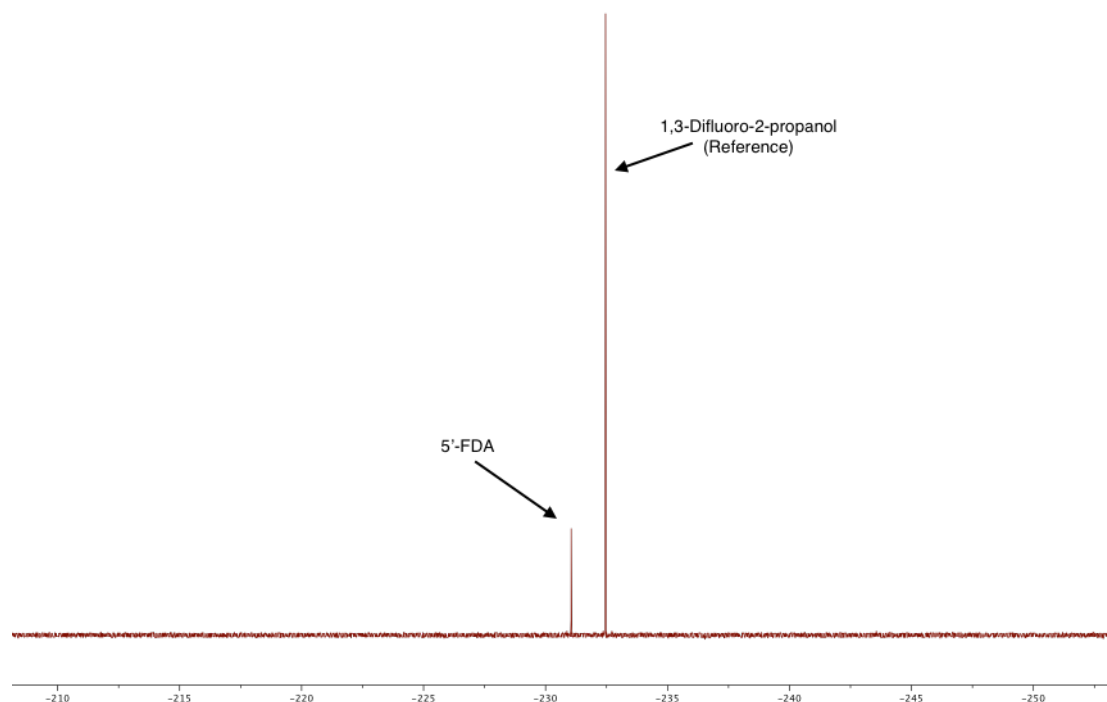
**Fig. 4**  $^{19}\text{F}$   $\{^1\text{H}\}$  NMR spectrum of the four-enzyme biotransformation of fluoride to 5-FHPA **2**, carried out as a one-pot experiment

### 3.3.2 Individual enzymatic analysis of the biotransformation from fluoride to 5-FHPA (**2**)

In addition to the control experiments carried out in the previous section, each step of the biotransformation was explored individually, to establish that the enzymes are fully functional. A study of each on the individual steps was carried out to deconstruct the overall biotransformation.

#### 3.3.2.1 *The fluorinase transformation*

The fluorinase was incubated with SAM and KF in phosphate-citrate buffer, (pH = 7.2, at 37 °C, 24 h). The outcome is shown in **Fig.5**. The only product is 5'-FDA, with added 1,3-difluoro-2-propanol as a reference.

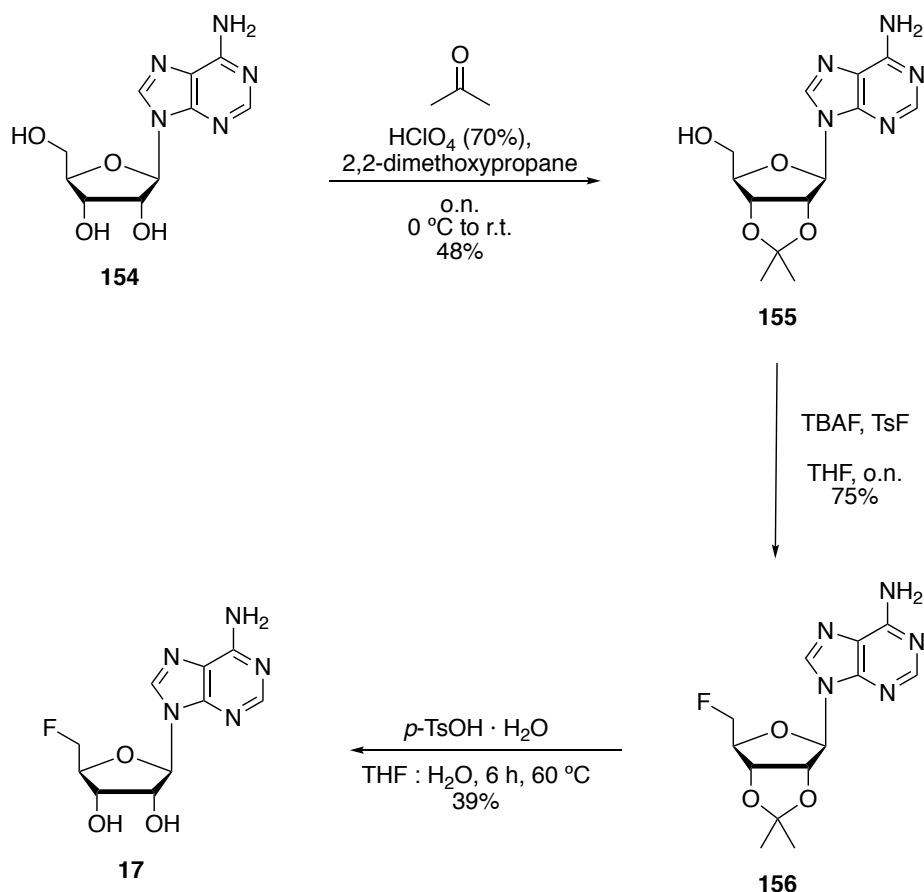


**Fig. 5**  $^{19}\text{F}$   $\{^1\text{H}\}$  NMR spectrum of the fluorinase assay

Although the clean production of 5'-FDA is established in terms of fluorinated products, the actual conversion from SAM cannot be established by  $^{19}\text{F}$  NMR spectroscopy.

### 3.3.2.2 *The purine nucleoside phosphorylase transformation*

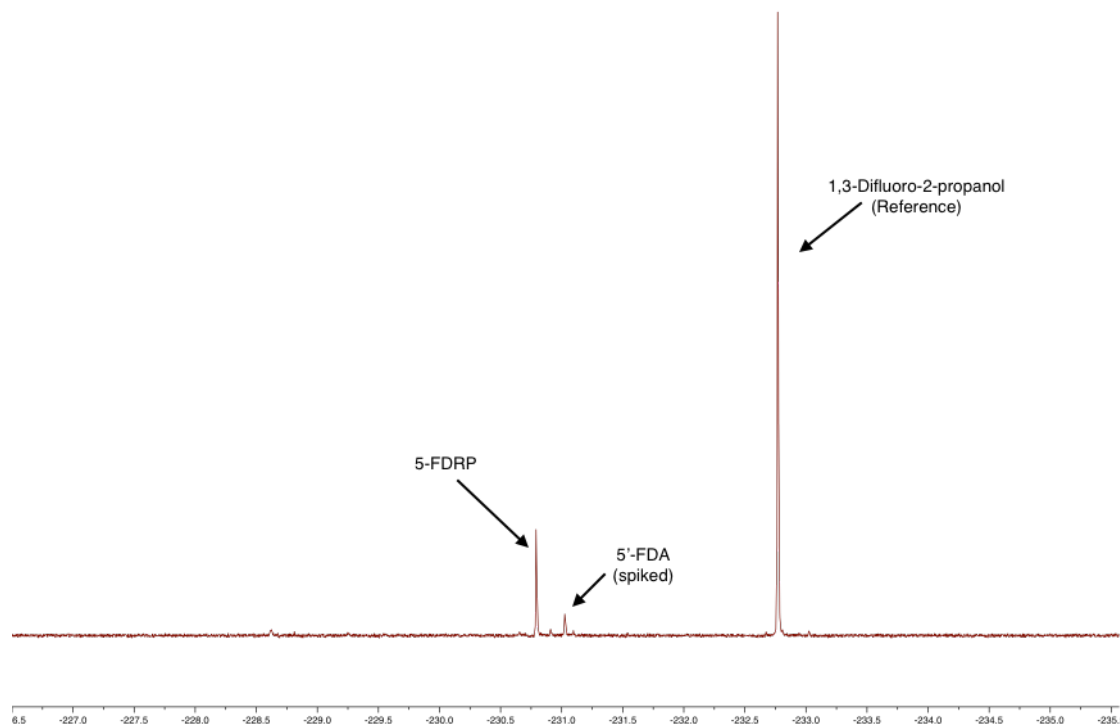
In the second step of the biotransformation, purine nucleoside phosphorylase (PNP) transforms 5'-FDA into 5-FDRP. 5'-FDA is not commercially available and was required to be synthesised from adenosine **154** by the three-step route illustrated in **Scheme 8**.<sup>11</sup>



**Scheme 8.** Synthetic route to 5'-FDA<sup>11</sup>

The synthetic route started with acetonide protection of the ribose of the adenosine (**154**), a reaction that afforded **155** in 48% yield. Tosylation and fluorination at C-5' were carried out in one step, leading to **156** (75%). The final deprotection step proved challenging - and was only successful with *p*-TsOH under reflux. 5'-FDA **17** was afforded in 39% yield.

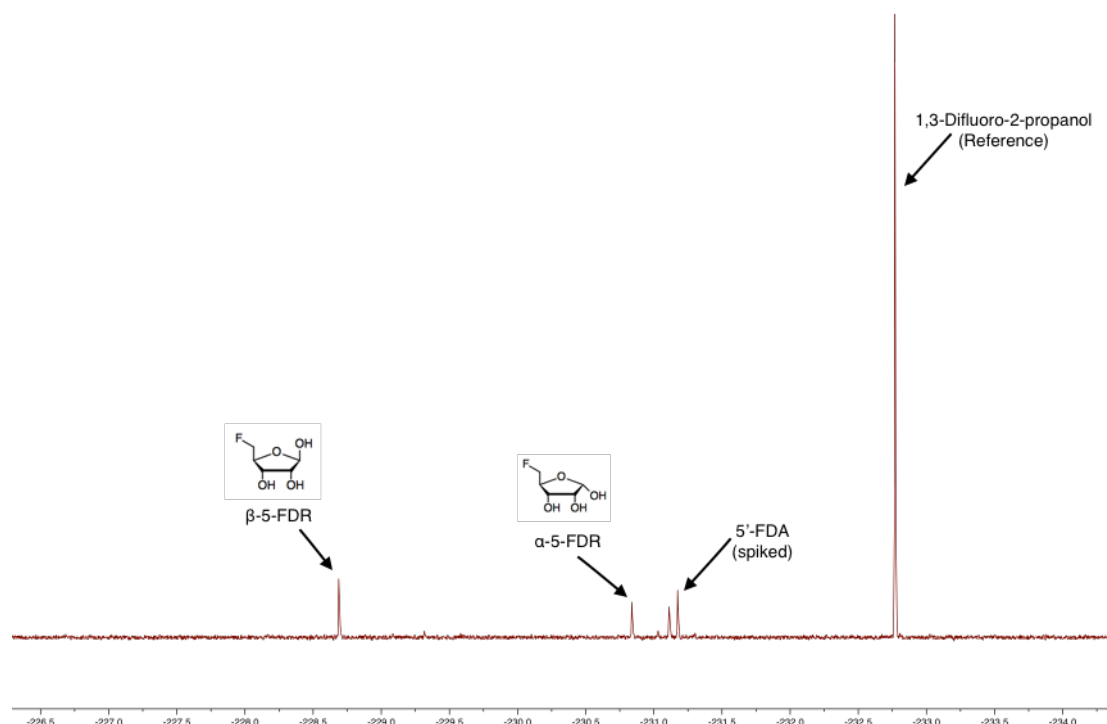
Incubation of synthetic 5'-FDA with the immobilised recombinant PNP (GSK) afforded complete conversion to 5-FDRP (**Fig. 6**). In order to prove the formation of 5-FDRP (given the very similar NMR shifts) it was necessary to spike with some 5'-FDA, which confirmed that the starting material was completely consumed and transformed into 5-FDRP (**Fig. 6**). The enzymatically-produced 5-FDRP was used as a substrate for the following transformation with the phytase.



**Fig. 6** PNP-catalysed transformation of 5'-FDA into 5-FDRP. 5'-FDA present in the spectrum was spiked post-incubation as proof of biotransformation

### 3.3.2.3 The phytase transformation

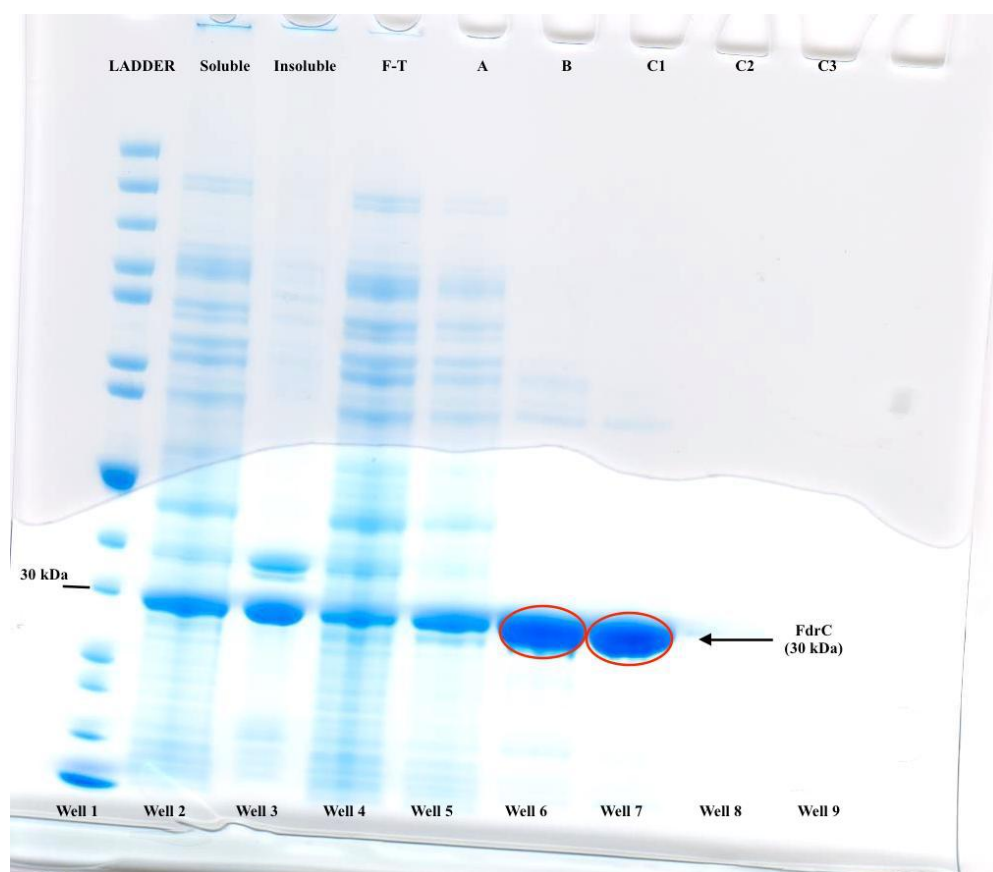
The third enzyme involved in the biotransformation is a dephosphorylase (FdrA). Given the lack of availability of this dephosphorylase from *S. sp.* MA37, it was substituted for a commercial phytase from wheat (Sigma-Aldrich). This phytase had the inconvenience of only being active under acidic conditions and at relatively high temperatures; therefore, this step had to be set up using appropriate conditions. 5-FDRP (obtained enzymatically, as discussed in **Section 3.3.2.2**, for the previous step) was incubated with the phytase in phosphate-citrate buffer (pH = 5.1), at 48 °C for 48 h. The  $^{19}\text{F}$   $\{^1\text{H}\}$  NMR spectrum of the reaction solution showed the formation of two new peaks, corresponding to the  $\alpha$  and  $\beta$  anomers of 5-FDR (Fig. 7).



**Fig. 7** Biotransformation of 5-FDRP into 5-FDR by phytase; residual 5'-FDA from spiking to the previous step can be observed

#### 3.3.2.4 The *FdrC* transformation

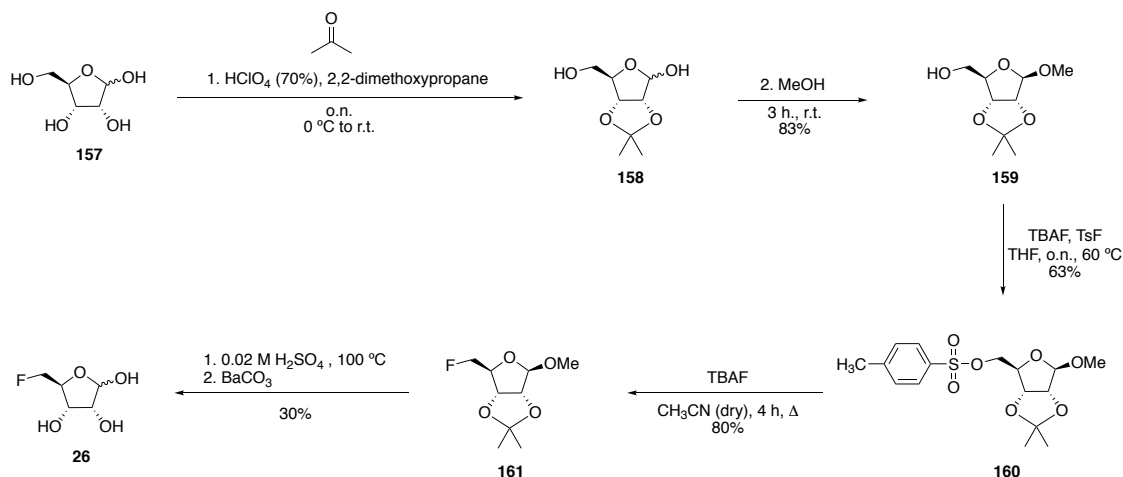
*FdrC* is an  $\text{NAD}^+$ -dependent short chain dehydrogenase, and is the enzyme responsible for the oxidation of 5-FDR to lactone **27**, and then hydrolysis to 5-FHPA **2**.<sup>1</sup> A synthetic plasmid with the *FdrC* gene was transformed into *E. coli* (BL21 gold), prior to over-expression by IPTG induction, and final purification on a Ni-NTA column, following the protocol<sup>1</sup> described in **Section 5.6.4** (Chapter 5). **Fig. 8** shows an image of the SDS-PAGE gel obtained from the purification of *FdrC*, which has a molecular weight of 30 kDa. Confirmation of the presence of *FdrC* was also obtained by MALDI mass spectrometry.



**Fig. 8** SDS-PAGE gel of over-expressed FdrC after purification on a Ni-NTA column. a) Well 1: Molecular weight reference ladder. b) Well 2: Soluble fraction. c) Well 3: Insoluble fraction. d) Well 4: Flow-through. e) Well 5: Wash buffer A, lysis buffer (20 mM Tris·HCl, 20 mM imidazole, 500 mM NaCl, pH=8). f) Well 6: Wash buffer B (20 mM Tris·HCl, 50 mM imidazole, 500 mM NaCl, pH=8). g) Well 7: Wash buffer C, elution buffer (20 mM Tris·HCl, 400 mM imidazole, 500 mM NaCl, pH=8) h) Well 8: Wash buffer C, elution buffer. i) Well 9: Wash buffer C, elution buffer

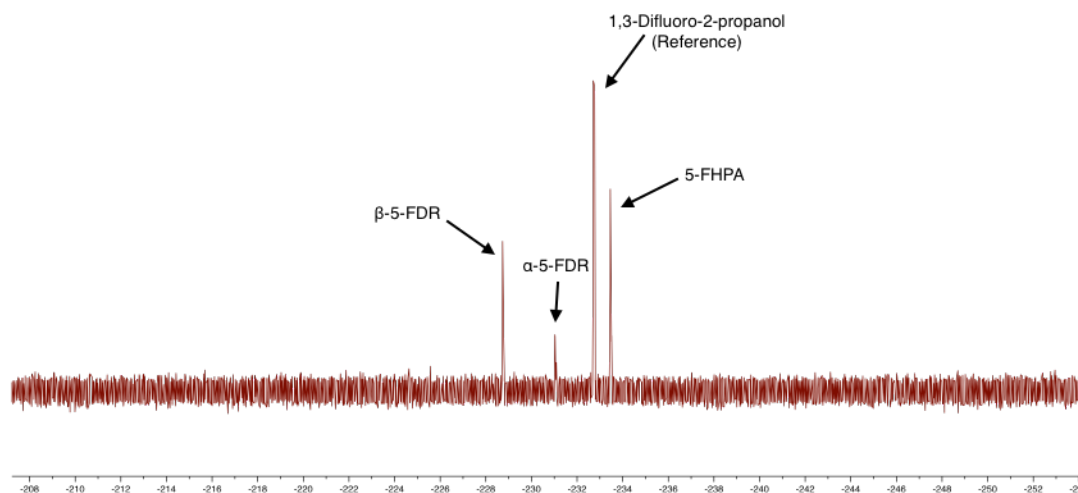
Since 5-FDR is not commercially available, it was required to be synthesised. The method used is reported in **Scheme 9**, and follows a protocol previously studied in the group.<sup>5,11,12</sup> The synthesis started with protection of ribose **157**, to generate acetonide **158**, followed by a methanolysis protection at the anomeric carbon to give **159**. Tosylation of the free C5 alcohol was followed by deoxofluorination to give acetonide **161**. Finally, removal of the acetonide under acidic conditions afforded 5-fluororibose (5-FDR, **26**). This final deprotecting step is challenging, and was only achieved in a 30% yield. The complications with the last step mainly reside in the fact that the reaction needs high temperature (>100 °C), but those high temperatures cause the sugar to decompose.





**Scheme 9.** Synthetic procedure for the chemical synthesis of 5-FDR (**26**)

For the enzymatic procedure, synthetic 5-FDR **26** was incubated with FdrC and  $\text{NAD}^+$  in phosphate-citrate buffer (pH = 7.2) at 37 °C for 24 h. It can be seen by  $^{19}\text{F}$   $\{^1\text{H}\}$  NMR (**Fig. 9**) that there is residual 5-FDR ( $\alpha$  and  $\beta$  anomers) in the extracts, and the reaction has gone to around 50% conversion.



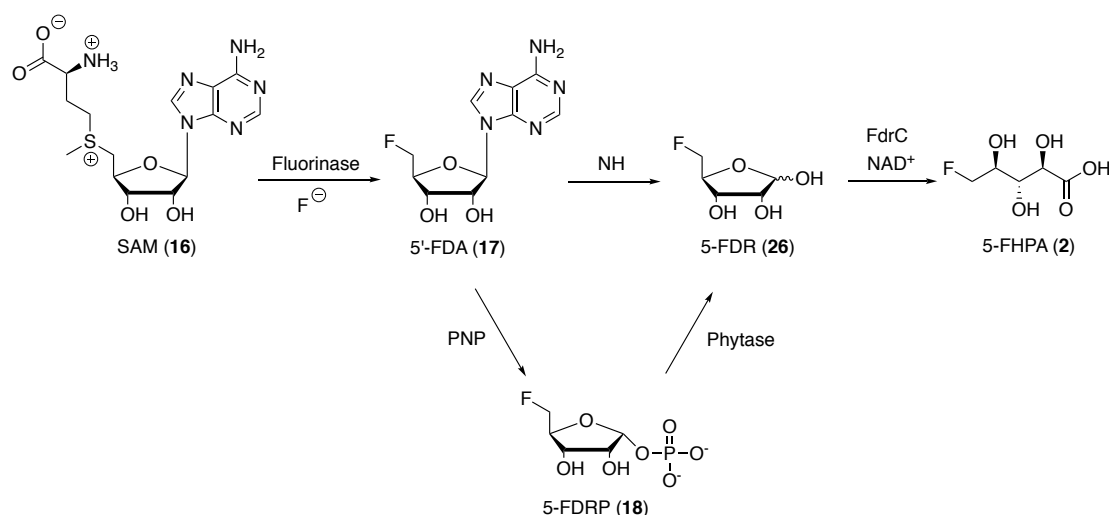
**Fig. 9** Biotransformation of 5-FDR **26** into 5-FHPA **2** with FdrC

### 3.4 Optimisation of the enzymatic route

#### 3.4.1 Using a nucleoside hydrolase

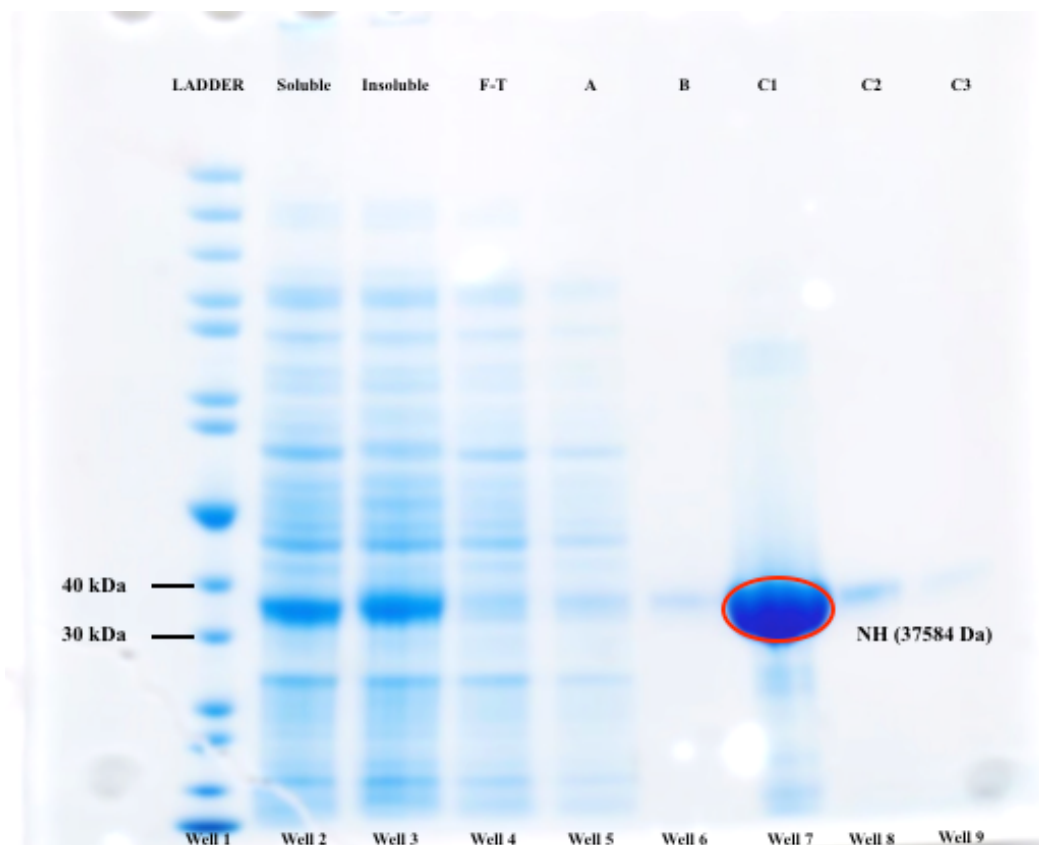
Both the individual experiments and the complete reconstitution from SAM and KF suggest that the phytase enzyme is a ‘bottle neck’, hindering the full conversion of fluoride to 5-FHPA **2**. Additionally, it is this step that forces a change of the pH and results into a more complex three-step procedure.

Therefore, the first optimisation of the route was aimed at avoiding the use of the phytase. A new biotransformation was envisaged, substituting the PNP and phytase for a nucleoside hydrolase (NH), an enzyme which can transform 5'-FDA directly into 5-FDR, in a one-step hydrolytic reaction,<sup>13</sup> as shown in **Scheme 10**.



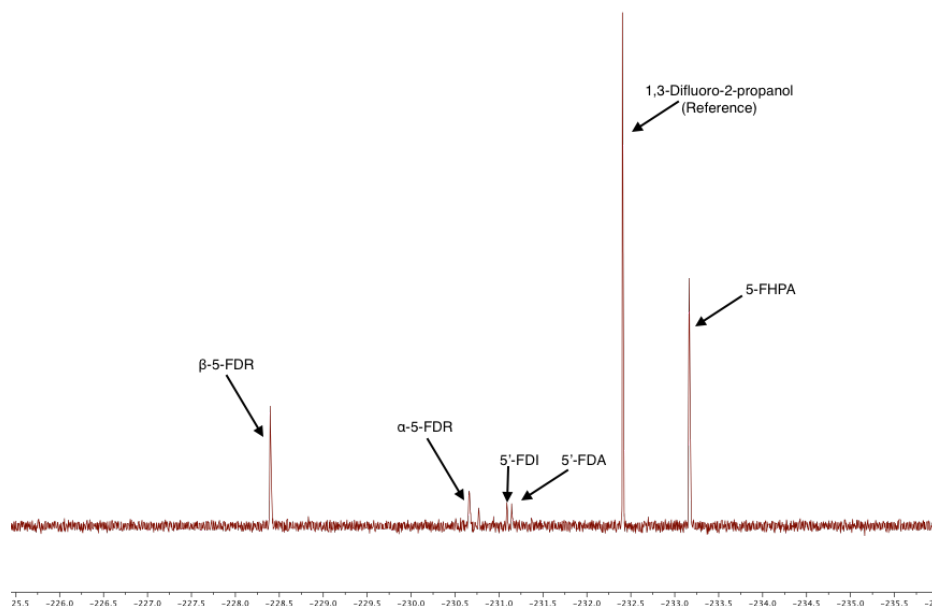
**Scheme 10.** Original biotransformation and alternative route using a NH

The nucleoside hydrolase from *Trypanosoma vivax* (an enzyme that was available in the lab) was over-expressed in *E. coli* and purified on a Ni-NTA column, following a previous protocol.<sup>11</sup> **Fig. 10** shows the resultant SDS-PAGE gel after purification. The NH, which has a molecular weight of approximately 37 kDa, appears below the 40 kDa mark. Confirmation of the presence of NH was also obtained by MALDI mass spectrometry. In order to preserve the stability and activity of the enzyme, it proved necessary to supplement the dialysis buffer with  $Ca^{2+}$  ions, since NH from *T. vivax* contains a calcium ion in its catalytic pocket, that participates in the binding to ribose.<sup>14</sup> In this case,  $CaCl_2$  was used.



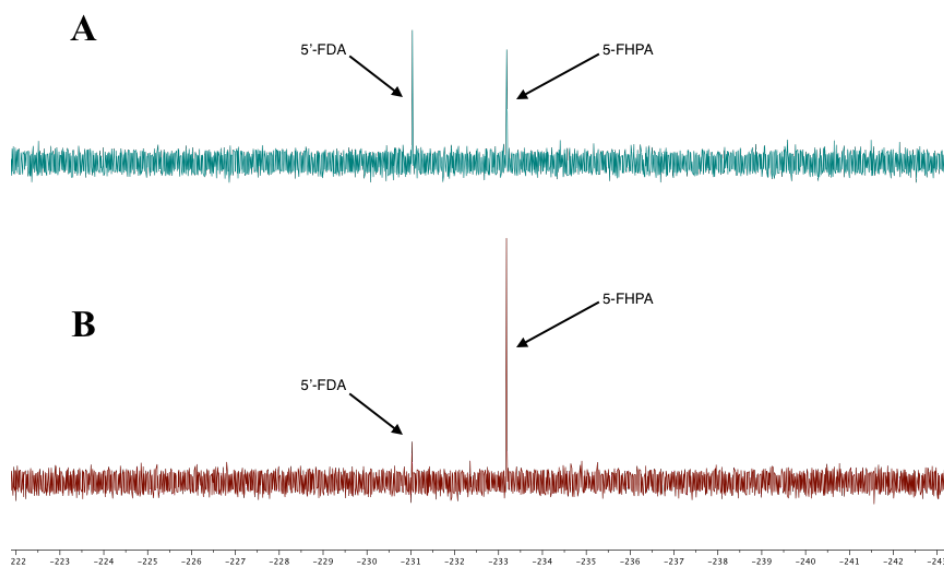
**Fig. 10** SDS-PAGE gel of the production of NH after a Ni-NTA purification column. a) Well 1: Molecular weight reference ladder. b) Well 2: Soluble fraction. c) Well 3: Insoluble fraction. d) Well 4: Flow-through. e) Well 5: Wash buffer A (10 mM phosphate buffer, 20 mM imidazole, 150 mM NaCl, pH = 7.8). f) Well 6: Wash buffer A. g) Well 7: Wash buffer C, elution buffer (10 phosphate buffer, 500 mM imidazole, 150 mM NaCl, pH = 7.8) h) Well 8: Wash buffer C, elution buffer. i) Well 9: Wash buffer C, elution buffer

The first experiments using the recombinant NH were carried out in PBS buffer, at a pH of 7.4. For this, and all the subsequent assays, the order of addition was as follows: SAM·2HCl, KF, the appropriate buffer, fluorinase, NH, FdrC and NAD<sup>+</sup>. Incubation was carried out at 37 °C, typically for 24 h. Under these conditions, the enzymatic reaction was significantly improved, and a complete bioconversion of fluoride to 5'-FHPA **2** was achieved. However, in the product mixture there was also residual 5'-FDA and 5-FDR, as well as the shunt product 5'-FDI, as illustrated in **Fig. 11**.



**Fig. 11**  $^{19}\text{F}\{^1\text{H}\}$  NMR spectrum of the production of 5-FHPA from SAM and fluoride using NH, instead of the PNP-phytase protocol

A new assay was carried out in phosphate buffer (pH = 8.0), with supplementation with  $\text{MgCl}_2$ .  $^{19}\text{F}\{^1\text{H}\}$  NMR spectroscopy showed 5-FHPA, but also some residual 5'-FDA (**Fig. 12**, upper). In an attempt to improve the production of 5-FHPA, the same experiment was repeated with an increased amount (4 times) of NH (**Fig. 12**, lower). This led to a more successful biotransformation.



**Fig. 12**  $^{19}\text{F}\{^1\text{H}\}$  NMR of 5-FHPA production from SAM and fluoride ion, carried out in phosphate buffer. A) 3 mg/mL of NH B) 12 mg/mL of NH

**Table 1** summarises all the conditions used, along with the outcome of each experiment. The most efficient conversion used phosphate buffer and NH concentration of 12 mg/mL (Entry 4).

**Table 1.** Summary of all conditions used for the biotransformation of fluoride to 5-FHPA

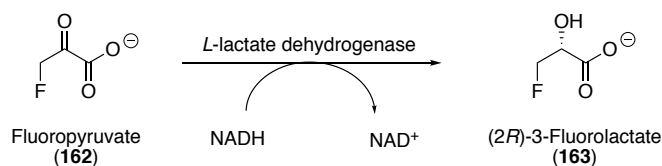
Entry	Buffer	pH	Mg <sup>2+</sup>	Incubation Time	NH Conc.	Outcome
1	Tris·HCl	7.8	+	24 h	3 mg/mL	5'-FDA:5-FHPA 1:1.5
2	Tris·HCl	7.8	+	24 h	6 mg/mL	5'-FDA:5-FHPA 2:1
3	Phosphate	8.0	+	24 h	3 mg/mL	5'-FDA:5-FHPA 1.1:1
4	Phosphate	8.0	+	24 h	12 mg/mL	5'-FDA:5-FHPA 1:7
5	PBS	7.2	-	48 h	3 mg/mL	<b>Fig. 11</b>

### 3.4.2 Implementation of an NAD<sup>+</sup> recycling protocol

Since the cofactor NAD<sup>+</sup> is consumed in the final enzymatic transformation, and it requires to be added in a stoichiometric amount, an NAD<sup>+</sup> recycling system was considered in order to regenerate NAD<sup>+</sup> from NADH.

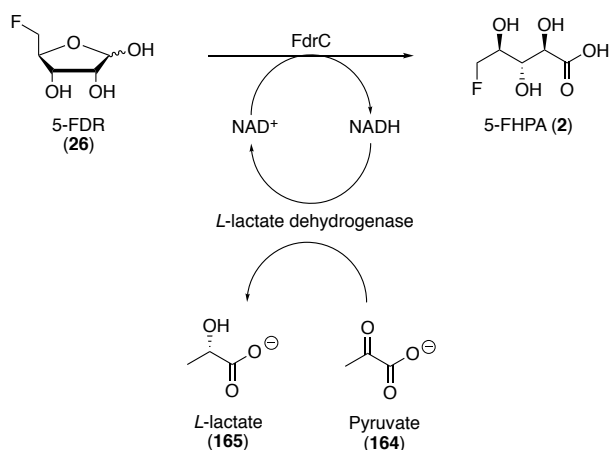
A traditional way to convert NADH into NAD<sup>+</sup> is to use lactate dehydrogenase (LDH) in the transformation of pyruvate into *L*-lactate.<sup>15</sup> A *L*-lactate dehydrogenase from rabbit muscle (Sigma Aldrich) was explored by incubation with fluoropyruvate (**162**) and NADH (**Scheme 11**), a transformation that was readily followed by <sup>19</sup>F NMR

spectroscopy. The transformation was satisfactory, and showed full conversion into (2*R*)-3-fluorolactate (**163**) by  $^{19}\text{F}$  NMR spectroscopy.



**Scheme 11.** Reduction of 3-fluoropyruvate **162** by LDH

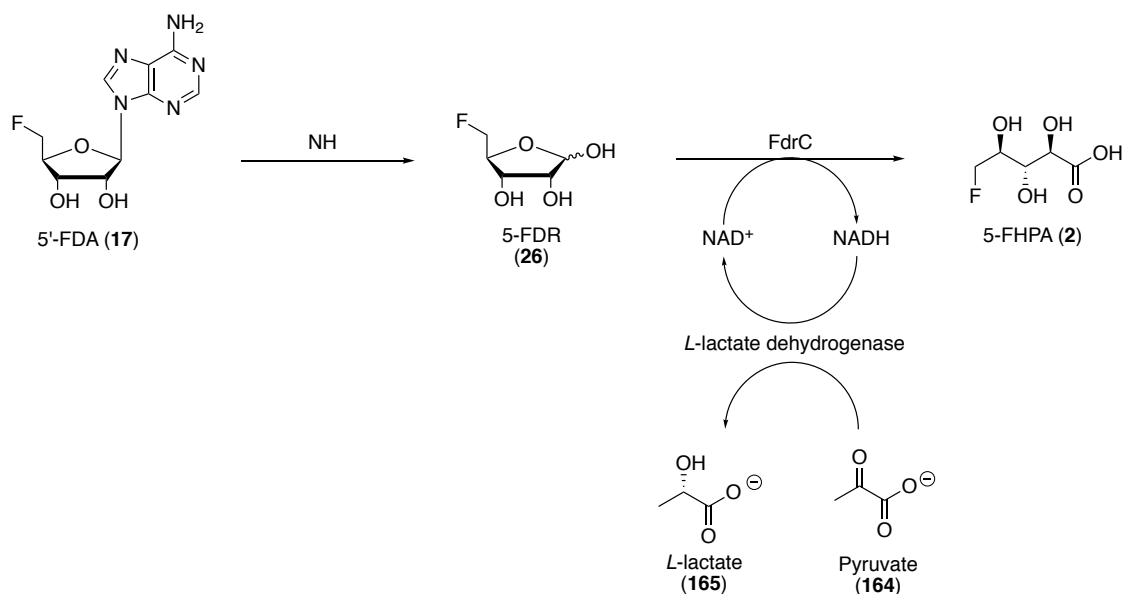
Given the successful generation of NAD<sup>+</sup> with LDH, this recycling system was explored for the last step of the biotransformation, as shown in **Scheme 12**.



**Scheme 12.** NAD<sup>+</sup> recycling system tested for the FdrC transformation of 5-FDR into 5-FHPA

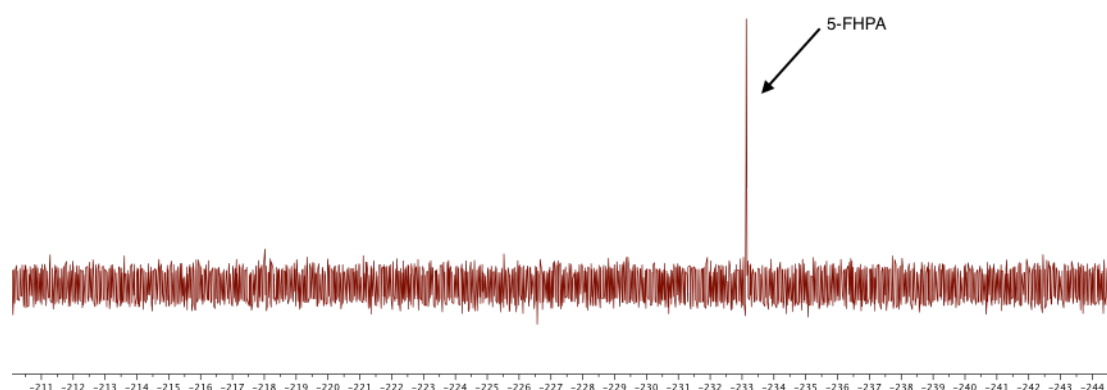
The experiment was explored in three different buffers to establish optimal conditions. These were phosphate (pH = 7.8); PBS (pH = 7.4); and Tris·HCl buffer (pH = 7.8). All three buffers showed conversion into 5-FHPA, indicating the regeneration cycle is active; however, residual 5-FDR is obvious in all three. Given that none of the experiments carried out with 5-FDR as starting material for the FdrC transformation into 5-FHPA showed full conversion, it is indicative that the enzymatic turnover is slow, perhaps with substrate or product inhibition.

In an attempt to avoid the saturation of FdrC, a two-step biotransformation starting from 5'-FDA was explored. This experiment would allow the gradual generation of 5-FDR, and perhaps improve the overall transformation. The assay was carried out in Tris·HCl buffer (pH = 7.5), at 37 °C over 20 h, and the outcome is shown in **Scheme 13**.



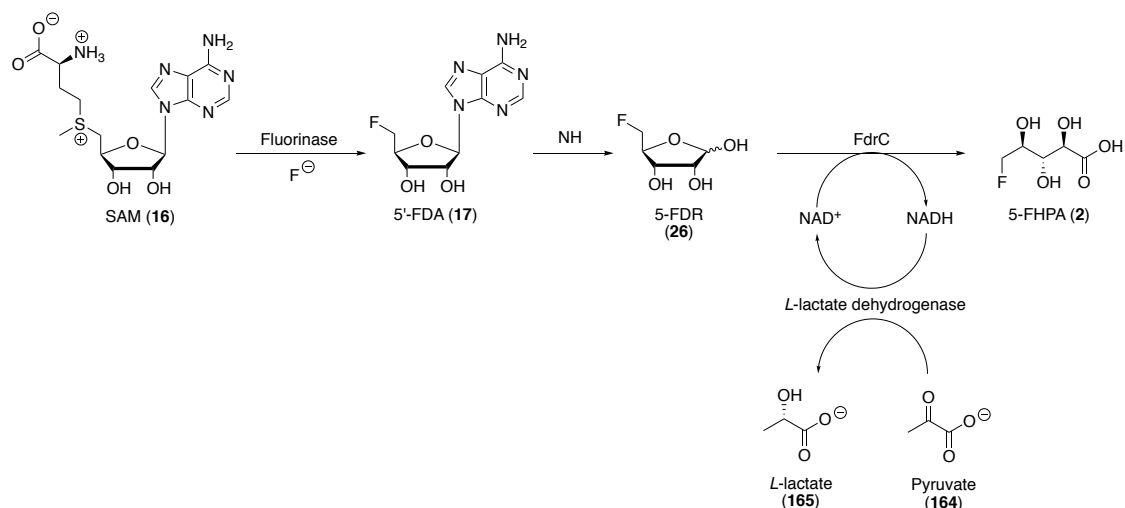
**Scheme 13.** Assay from 5'-FDA to 5-FHPA, with LDH as  $\text{NAD}^+$  regenerating agent

The transformation was successful, leading to full conversion into 5-FHPA (**Fig. 13**), suggesting that a slower *in situ* generation of 5-FDR improves the overall conversion. This outcome suggests that 5'-FDA (**17**) is more suitable than 5-FDR (**26**) as a starting point for this biotransformation to 5-FHPA (**2**).



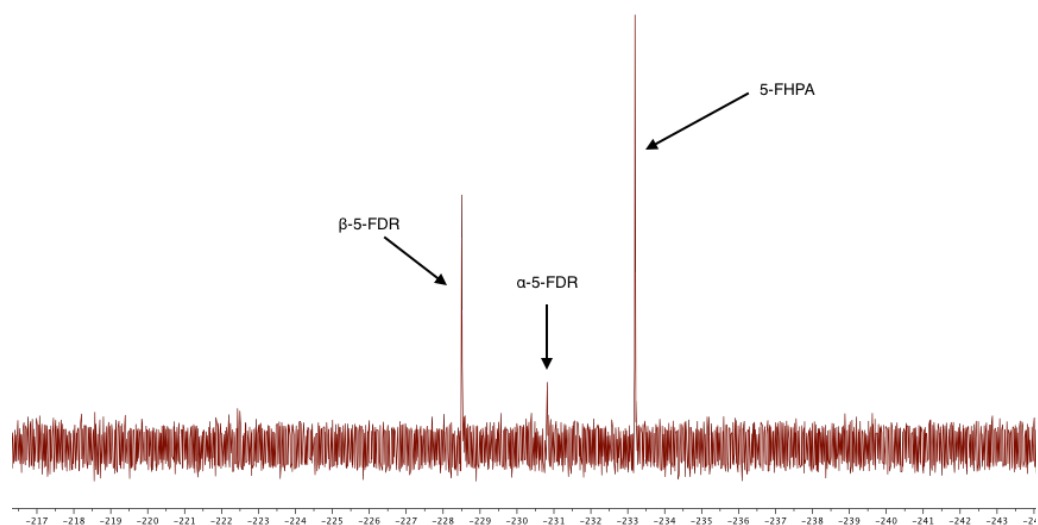
**Fig. 13**  $^{19}\text{F}$   $\{^1\text{H}\}$  NMR of 5'-FDA into 5-FHPA with an LDH recycling system of  $\text{NAD}^+$

Finally, the 'complete' transformation of fluoride into 5-FHPA with LDH as a recycling agent for  $\text{NAD}^+$  was explored (**Scheme 14**).



**Scheme 14.** Enzymatic transformation of inorganic fluoride into 5-FHPA, with NAD<sup>+</sup> regeneration

The experiment was run in Tris·HCl buffer (pH = 7.5) and with Mg<sup>2+</sup> supplementation. Incubation was carried out at 37 °C overnight (20 h). The outcome is shown by <sup>19</sup>F NMR spectroscopy in **Fig. 14**.



**Fig. 14** <sup>19</sup>F{<sup>1</sup>H} NMR spectrum of the generation of 5-FHPA from inorganic fluoride and SAM, with LDH for NAD<sup>+</sup> recycling

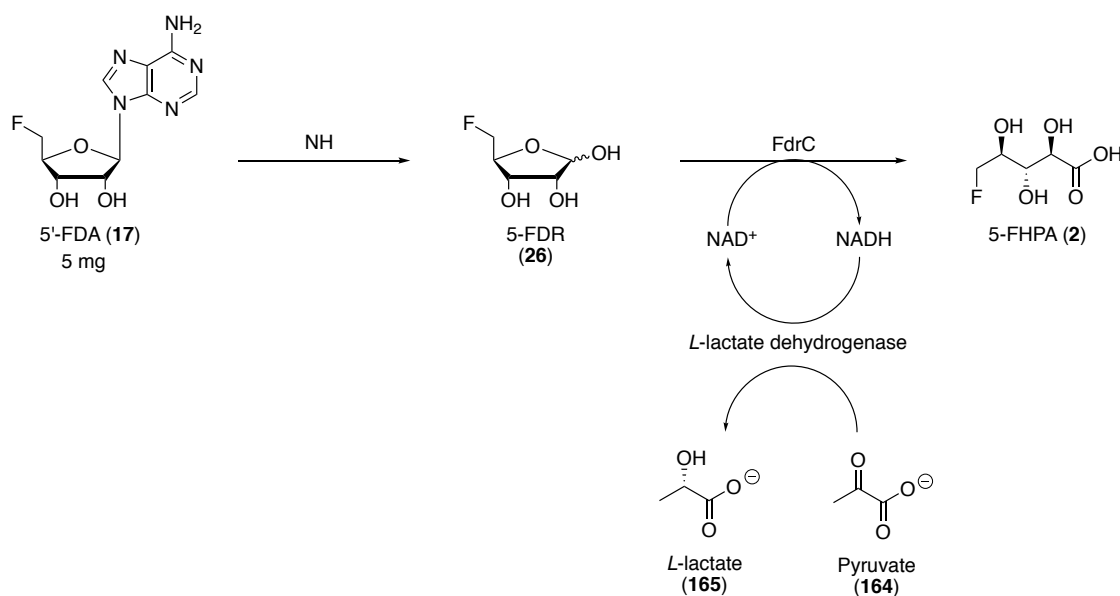
This experiment starts from inorganic fluoride, and involves four enzymes, to generate 5-FHPA. However, it was obvious that 5-FDR (α and β anomers) accumulated in the reaction, suggesting again a sluggish conversion of 5-FDR into 5-FHPA.



### 3.5 Scale-up of the 5-FHPA biotransformation

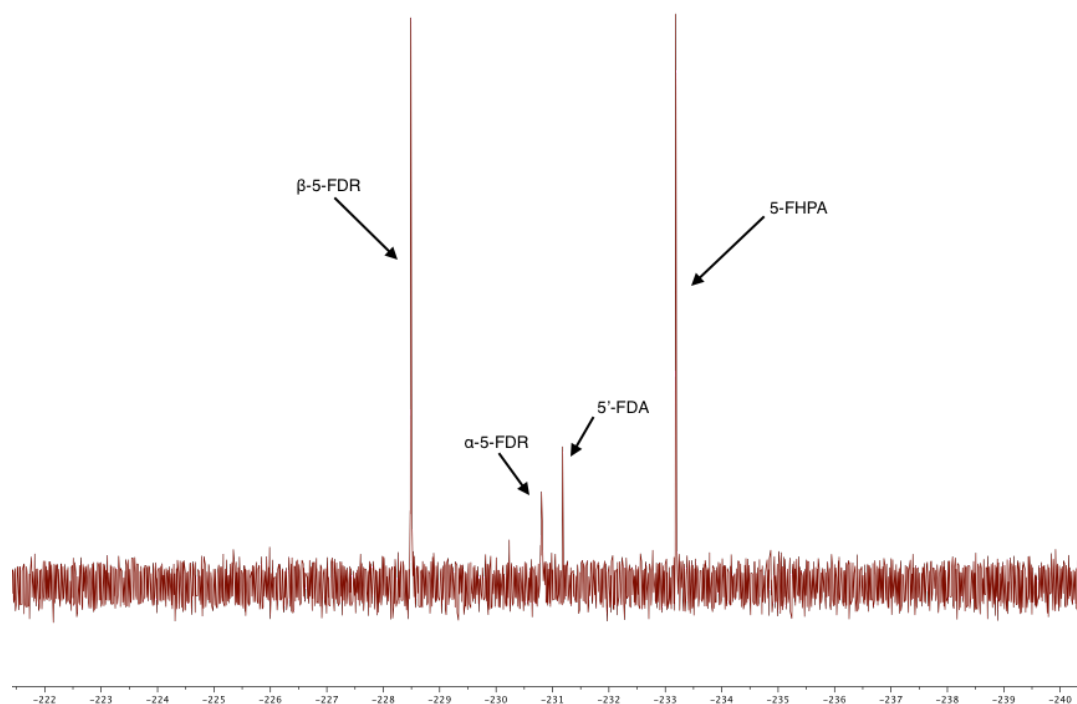
Up to this point all of the experiments were carried out on an analytical scale. The next challenge in the optimisation process was to scale-up the production of 5-FHPA to milligram amounts.

Given that the most successful biotransformation of 5-FHPA so far started from 5'-FDA (17), initial optimisation explored the biotransformation from this metabolite (5 mg), using LDH as the regenerating agent of  $\text{NAD}^+$  (**Scheme 15**). Incubations were carried out in Tris·HCl buffer at a pH of 7.5, 37 °C over an extended incubation time (120 h).



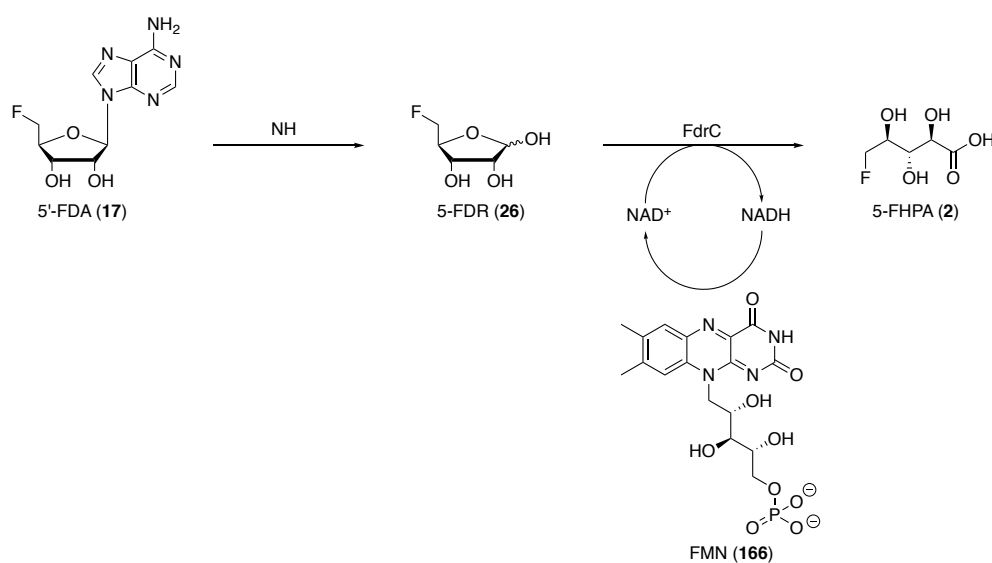
**Scheme 15.** Preparative scale (5 mg) of the biotransformation of 5'-FDA to 5-FHPA

Aliquots were analysed after 24 h, 54 h and 72 h, prior to final  $^{19}\text{F}$  NMR analysis after 120 h. The reaction had stopped after 24 h, when the experiment was essentially complete. The accumulation of 5-FDR anomers is observed again, along with a low level of residual 5'-FDA, as well as the product 5-FHPA (**Fig. 15**). The fact that 5-FDR is the major product indicates that the final FdrC step is not fully functional – possibly due to an incomplete recycling of  $\text{NAD}^+$ , or perhaps due to substrate inhibition.



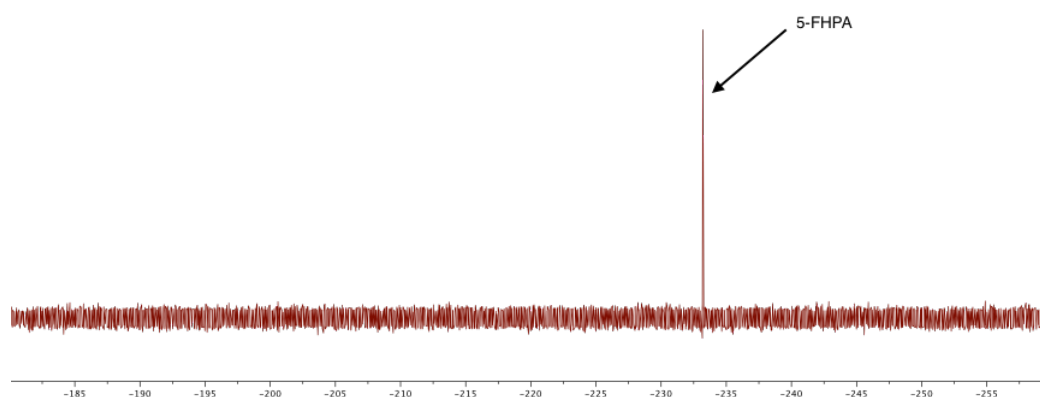
**Fig. 15**  $^{19}\text{F}\{^1\text{H}\}$  NMR spectrum of one of the 5 mg scale-ups after 120 h; residual 5-FDR is the major metabolite

In view of the previous result, a new recycling system for  $\text{NAD}^+$  was explored using flavin mononucleotide (FMN, **166**).<sup>16</sup> FMN acts in conjugation with NADH, oxidising the cofactor back to  $\text{NAD}^+$ , and generating  $\text{FMNH}_2$  by hydride transfer in a non-enzymatic reaction.<sup>16</sup> FMN was first tested on an analytical scale (**Scheme 16**).



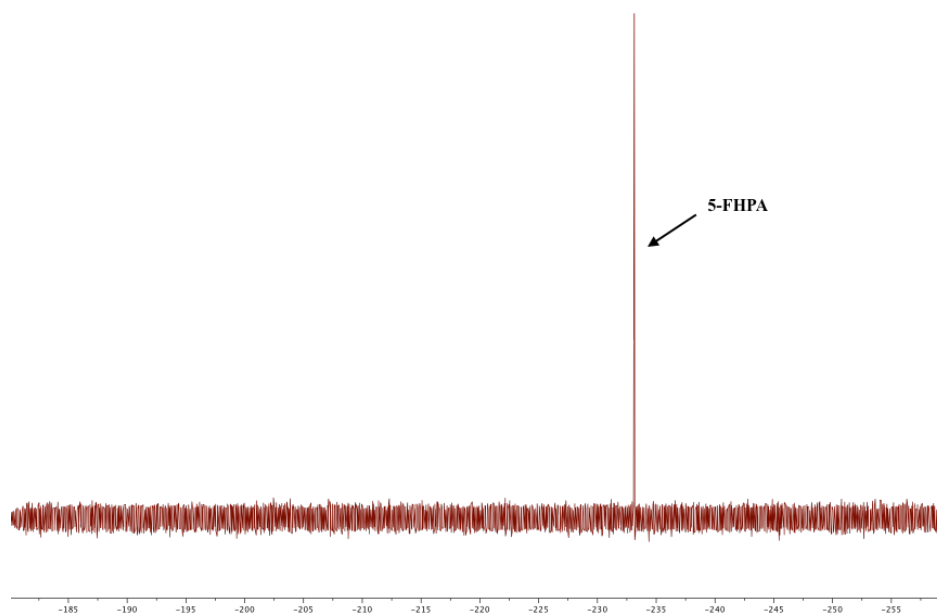
**Scheme 16.** Assay using FMN as a recycling agent for  $\text{NAD}^+$

The experiment described in **Scheme 16** was carried out in Tris·HCl buffer (pH = 7.5), at 37 °C for 20 h.  $^{19}\text{F}$  NMR spectroscopy of the product shows the sole formation of 5-FHPA, indicating FMN is highly efficient. The  $^{19}\text{F}$   $\{^1\text{H}\}$  NMR spectrum is shown in **Fig. 16**.



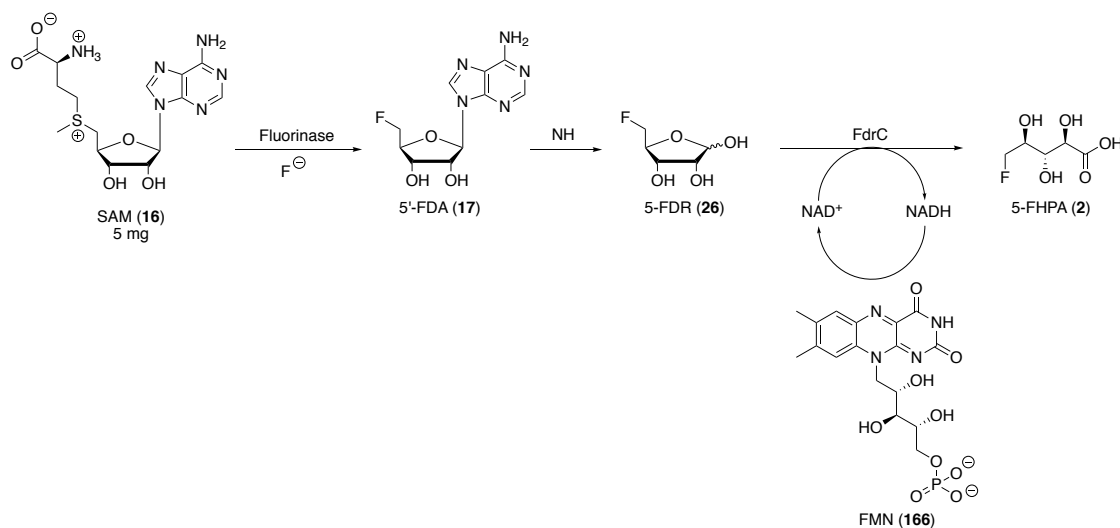
**Fig. 16**  $^{19}\text{F}$   $\{^1\text{H}\}$  NMR shows conversion of 5'-FDA to 5-FHPA by FdrC, with FMN/NAD<sup>+</sup> cofactor recycling

After this transformation, a scale-up of the procedure was carried out using 5'-FDA (5 mg), following the method described in above. After 24 h at 37 °C, the only product that could be observed by  $^{19}\text{F}$   $\{^1\text{H}\}$  NMR was 5'-FHPA (**Fig. 17**), indicating a complete conversion of 5'-FDA to 5-FHPA.



**Fig. 17** Production of 5-FHPA from 5'-FDA (5 mg), with FMN as recycling agent

Encouraged by this result, a ‘complete’ biotransformation of 5-FHPA from fluoride was explored, as illustrated in **Scheme 17**, starting in this case with SAM (5 mg).



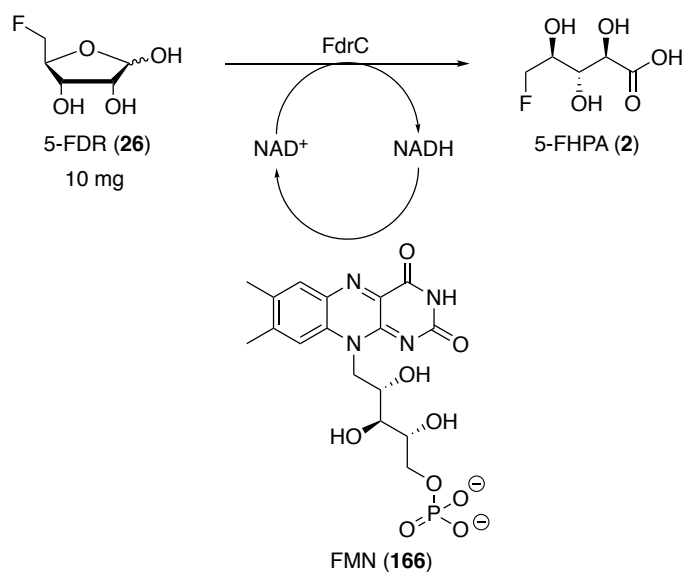
**Scheme 17.** Biotransformation from SAM and KF to 5-FHPA with FMN to recycle  $NAD^+$

After a 24 h incubation, the only observed peak – apart from that of fluoride – was 5'-FDA, thus the biotransformation appeared to stop at the NH step. In order to explore what might be the cause of this arrest, some control assays were carried out.

An analytical reaction with 5'-FDA and the nucleoside hydrolase was conducted, to check whether the enzyme had lost its activity, however, there was again a full conversion of 5'-FDA into 5-FDR, indicating that the enzyme was still active.

Another possibility was that the amount of SAM added (5 mg) was excessive, and perhaps downregulated the pathway in some way. To test this possibility, an analytical scale experiment from fluoride (**Scheme 17**) was carried out. This resulted in the formation of some 5-FHPA, but also in residual 5'-FDA and 5-FDR. Therefore, it was established again that SAM and inorganic fluoride do not serve as efficient starting materials for this biotransformation.

A scale-up reaction from 5-FDR (10 mg) was carried out to study whether the single FMN step would be more efficient than the two-step experiment starting from 5'-FDA. Accordingly, 5-FDR was incubated in Tris·HCl buffer at 37 °C and for 24 h (pH = 7.5), in the procedure illustrated in **Scheme 18**.



**Scheme 18.** Biotransformation of 5-FDR to 5-FHPA

The experiment was successful, but only progressed to about 25% conversion. 5-FDR remained in a ratio 3:1 with 5-FHPA. Once again, 5'-FDA proved to be the most efficient starting material for this enzymatic synthesis. Therefore, further experiments focused solely on this as a starting point for the biotransformation.

Fresh batches of enzymes (NH and FdrC) were prepared before increasing the amount of starting material, to avoid the loss of activity over time. An assay was set up with 50 mg of 5'-FDA, following the protocol illustrated in **Scheme 16** (page 158). The experiment was incubated at 37 °C, and monitored regularly over one week.

Fluorine NMR analysis after 20 h of incubation showed formation of 5-FHPA, with residual amounts of 5'-FDA and 5-FDR. Integration of the peaks indicated that the ratio of 5'-FDA:5-FDR:5-FHPA was 1.3:1:1.3. After 48 h, the amount of 5'-FDA had considerably decreased, and 5-FHPA was already the major product in the mixture. At day four, the ratio 5'-FDA:5-FDR:5-FHPA was 0.1:1:1.7. This indicated that although the enzymes were still active, FdrC activity was decreasing, and lowering the overall conversion. This allowed the accumulation of 5-FDR in the mixture. One week after the start of the assay, the enzymes were no longer active, but the conversion into 5-FHPA was improved from day 4. **Fig. 18** shows the evolution of this experiment, at days 2 and 4. Notably, no 5'-FDA was left at day 4, and the ratio of 5-FDR to 5-FHPA was 1:2, as illustrated in **Fig. 19**.

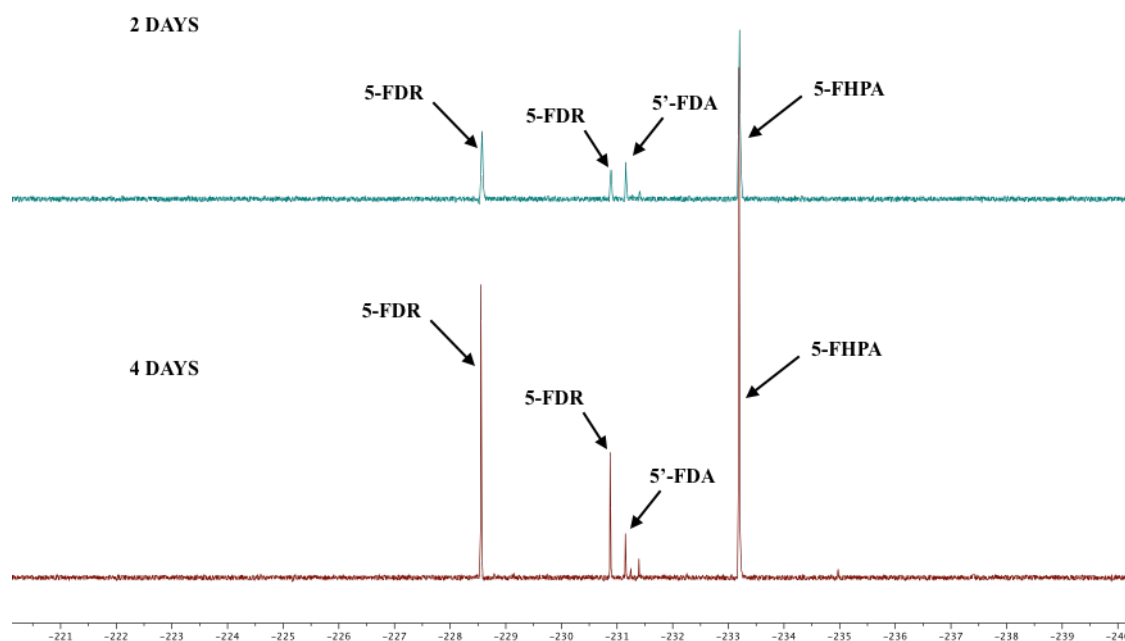


Fig. 18 Scale up from 5'-FDA (50 mg), at day 2 and day 4

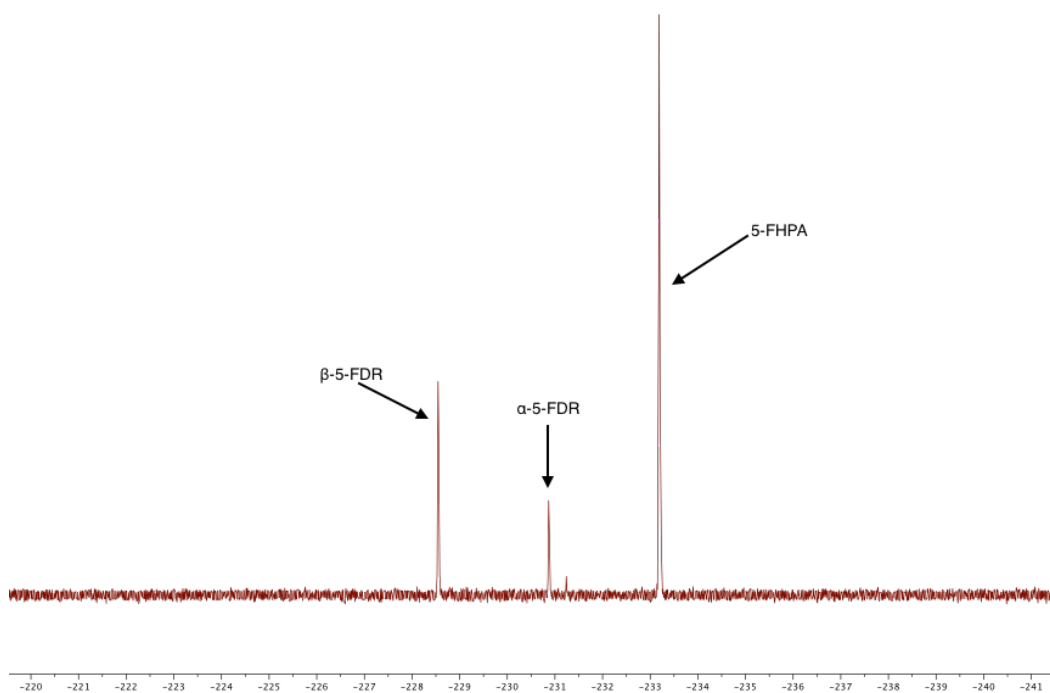


Fig. 19  $^{19}\text{F}$   $\{^1\text{H}\}$  NMR of the scale-up biotransformation of 5'-FDA into 5-FHPA, after one week

At this point, all reasonable optimisations had been implemented; such as starting with 5'-FDA, FMN as a functional recycling system for NAD<sup>+</sup>, freshly prepared enzymes, and the effect of different buffers, pH conditions and supplements (Mg<sup>2+</sup>).

Enzymatic biotransformation using 5'-FDA (50 mg) successfully led to the formation of 5-FHPA in a 60% conversion, as determined by <sup>19</sup>F NMR spectroscopy. Purification of a sample of the final product (5-FHPA, **2**) proved to be extremely challenging, due to its very high water solubility, and no chromophore characteristics for HPLC detection. This remains to be optimised.

### 3.6 Conclusions

The *in vitro* reconstitution of the biosynthetic pathway of 5-FHPA from fluoride was successfully achieved. The procedure started from SAM and inorganic fluoride, which were incubated with recombinant fluorinase from *S. cattleya*, recombinant PNP immobilised in Sepabeads®, a commercial phytase from wheat and the recombinant FdrC from a codon-optimised synthetic gene, along with  $\text{NAD}^+$ . Due to the optimal pH for the phytase being significantly lower than that from the other enzymes, the process was carried out as a one-pot, three-step method; which involved two changes of pH.

Optimal conditions for the combined chemical and enzymatic production of 5-FHPA involved 5'-FDA as a starting material for a two-step enzymatic synthesis that includes FMN as recycling agent for  $\text{NAD}^+$ .

Biotransformations from SAM and fluoride or from synthetic 5-FDR, did not result in a full conversion into 5-FHPA.

Finally, a scale-up experiment for the enzymatic production of 5-FHPA was successfully carried out from 50 mg of 5'-FDA, in a 60% conversion. In this biotransformation, the enzymes NH and FdrC were used, therefore avoiding any change of pH. Added FMN enabled the *in situ* recycling of  $\text{NAD}^+$  at a catalytic level.



### 3.7 References

- 
- <sup>1</sup> L. Ma, A. Bartholomé, M. H. Tong, Z. Qin, T. Shepherd, K. Kyeremeh, H. Deng, D. O'Hagan; *Chem. Sci.* **2015**, *6*, 1414.
- <sup>2</sup> H. Deng, L. Ma, N. Bandaranayaka, Z. Qin, G. Mann, K. Kyeremeh, Y. Yu, T. Shepherd, J. H. Naismith, D. O'Hagan; *ChemBioChem*, **2014**, *15*, 364.
- <sup>3</sup> S. Huang, L. Ma, M. H. Tong, Y. Yu, D. O'Hagan, H. Deng; *Org. Biol. Chem.* **2014**, *12*, 4828.
- <sup>4</sup> D. O'Hagan, C. Schaffrath, S. L. Cobb, J. T. G. Hamilton, C. D. Murphy; *Nature*, **2002**, *416*, 279.
- <sup>5</sup> S. L. Cobb, H. Deng, J. T. G. Hamilton, R. P. McGlinchey, D. O'Hagan; *Chem. Commun.* **2004**, *0*, 592.
- <sup>6</sup> a) M. J. Pugmire, S. E. Ealick; *Biochem J.* **2002**, *361*, 1. b) C. Mao, W. J. Cook, M. Zhou, A. A. Federov, S. C. Almo, S. E. Ealick; *Biochem.* **1998**, *37*, 7135.
- <sup>7</sup> M. Sanada, T. Miyano, S. Iwadare, J. M. Williamson, B. H. Arison, J. L. Smith, A. W. Douglas, J. M. Liesch, E. Inamine; *Antibiotics*, **1986**, *39*, 259.
- <sup>8</sup> A. S. Eustáquio, R. McGlinchey, Y. Liu, C. Hazzard, L. L. Beer, G. Florova, M. M. Alhamadsheh, A. Lechner, A. J. Kale, Y. Kobayashi, K. A. Reynolds, B. S. Moore; *PNAS*, **2009**, *106*, 12295.
- <sup>9</sup> H. Deng, S. M. Cross, R. P. McGlinchey, J. T. G. Hamilton, D. O'Hagan; *Chem. Biol.* **2008**, *15*, 1268.
- <sup>10</sup> F. G. Peers; *Biochem. J.* **1953**, *53*, 102.
- <sup>11</sup> M. Onega, PhD Thesis (University of St Andrews), **2009**.
- <sup>12</sup> X.-G. Li, S. Dall'Angelo, L. F. Schweiger, M. Zanda, D. O'Hagan; *Chem. Commun.* **2012**, *48*, 5247.
- <sup>13</sup> M. Onega, J. Domarkas, H. Deng, L. F. Schweiger, T. A. D. Smith, A. E. Welch, C. Plisson, A. D. Gee, D. O'Hagan; *Chem. Commun.* **2010**, *46*, 139.

---

<sup>14</sup> W. Versées, K. Decanniere, R. Pellé, J. Depoorter, E. Brosens, D. W. Parkin, J. Steyaert; *J. Mol. Biol.* **2001**, 307, 1363.

<sup>15</sup> L. P. B. Gonçalves, O. A. C. Antunes, G. F. Pinto, E. G. Oestreicher; *J. Mol. Catal. B Enzym.* **1998**, 4, 67.

<sup>16</sup> J. B. Jones, K. E. Taylor; *Can. J. Chem.* **1976**, 54, 2969.

## Chapter 4. Exploring the fluorinase pathway of *Streptomyces* sp. MA37

### 4.1 Introduction

Studying the biosynthesis of organofluorine compounds is a field of interest, given the scarcity of fluorinated natural products on Earth, and the potential to make organofluorines by biotransformations from fluoride. Only six fluorinated natural products are known.<sup>1</sup> The most recent was isolated and characterised in 2014 from the bacterium *Streptomyces* sp. MA37, and was identified as (2*R*,3*S*,4*S*)-5-fluoro-2,3,4-trihydroxypentanoic acid (5-FHPA, **2**). This product was biosynthesised along with fluoroacetate (FAC, **1**) and 4-fluorothreonine (4-FT, **4**).<sup>1</sup> Although 5-FHPA is the only novel fluorometabolite identified from *S.* sp MA37 so far, several additional fluorine peaks have been observed (by <sup>19</sup>F NMR spectroscopy) in profile studies of the organism.

*Streptomyces* sp. MA37 was collected in a fluoride rich soil in Ghana.<sup>1</sup> Sequencing of its full genome led to the identification of *flA1*, a fluorinase gene with 87% homology of that of the original fluorinase (*flA*) from *S. cattleya*.<sup>1</sup>

The original *fl* gene locus (*S. cattleya*) was first characterised in 2006 by Spencer *et al.*,<sup>2</sup> now known as the ‘Spencer cluster’. Cloning and analysis of this DNA locus showed that it encoded several genes around the fluorinase (*flA*), that could potentially be involved in fluorometabolite biosynthesis.<sup>2</sup> The genome map for this cluster, as well as the amino acid length for each gene and its expected function is shown in **Fig. 1**.<sup>2</sup>

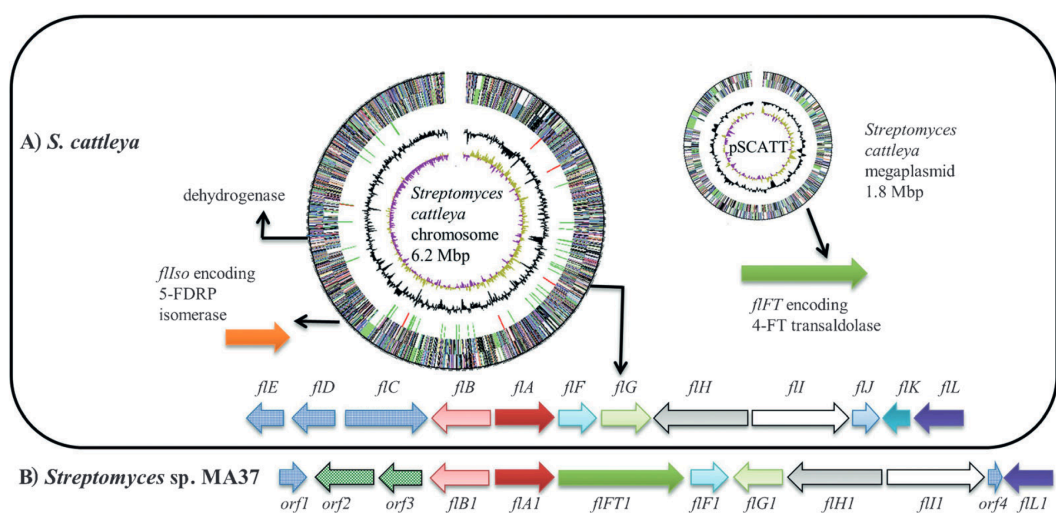
<i>flE</i>	<i>flD</i>	<i>flC</i>	<i>flB</i>	<i>flA</i>	<i>flF</i>	<i>flG</i>	<i>flH</i>	<i>flI</i>	<i>flJ</i>	<i>flK</i>	<i>flL</i>
←	←	→	←	→	→	←	←	→	→	←	←

ORF	Start/Stop (bp)	Length (aa)	Function/Homology
E	127-795c	222	DNA binding regulatory protein
D	854-1504c	216	Dehalogenase/Phosphatase
C	1845-3038	397	MFS permease
B	3053-3952c	299	5'-FDA phosphorylase
A	4172-5071	299	5'-FDA synthase
F	5196-5753	185	DNA binding regulatory protein
G	5947-6651c	234	DNA binding regulatory protein
H	6648-8051c	467	Na <sup>+</sup> /H <sup>+</sup> antiporter
I	8317-9786	489	Adenosylhomocysteine hydrolase
J	9802-10197	131	regulatory protein
K	10172-10591c	139	Fluoroacetyl-CoA thioesterase
L	10696-11373c	225	DNA binding regulatory protein

**Fig. 1** The Spencer Cluster, as published in 2006<sup>5</sup>

**Fig. 2** is an illustration of the comparison between the genome maps for *S. cattleya* and *S. sp. MA37*.<sup>2</sup>



**Fig. 2** Aligned genome maps for *S. cattleya* (the Spencer Cluster) and *S. sp. MA37*<sup>2</sup>

**Table 1** displays the anticipated functions for each of the open reading frames (ORF) within the fluorometabolite gene cluster of *S. sp.* MA37 and its homology to the original *S. cattleya*.<sup>2</sup>

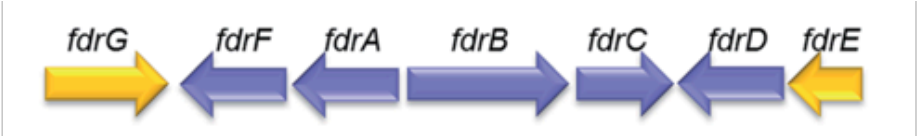
**Table 1.** Homology between the genomic functions of *S. sp.* MA37 and *S. cattleya*<sup>2</sup>

ORF	Expected Function	AA	Identity (%)	Homologue in <i>S. cattleya</i>
ORF1	Duf109	114	-	-
ORF2	Carboxylate/Amino Acid/ Amine Transporter	316	76%	SCAT_p0565
ORF3	YbaK/prolyl- <sup>t</sup> RNA synthetase	191	81%	SCAT_p0564
<i>flB1</i>	5'-FDA phosphorilase	299	66%	<i>flB</i>
<i>flA1</i>	5'-FDA synthase	299	88%	<i>flA</i>
<i>flFT1</i>	4-fluorothreonine transaldolase	625	72%	SCAT_p0562
<i>flF1</i>	DNA binding regulatory protein	171	72%	<i>flF</i>
<i>flG1</i>	DNA binding regulatory protein	226	61%	<i>flG</i>
<i>flH1</i>	Na <sup>+</sup> /H <sup>+</sup> antiporter	451	57%	<i>flH</i>
<i>flI1</i>	Adenosylhomocysteine hydrolase	487	82%	<i>flI</i>
ORF4	Duf109 fragment	94	-	-
<i>flL1</i>	DNA binding regulatory protein	224	57%	<i>flL</i>

This section of the genome is responsible for the formation of fluoroacetate (**1**) and 4-fluorothreonine (**4**) as observed in both *S. cattleya* and *S. sp. MA37*. However, the metabolic pathway in *S. sp. MA37* comprises an additional branch, which starts from 5-FDRP (**18**), and is expected to lead to the formation of 5-FHPA (**2**), as well as the additional unknown fluorometabolites.<sup>1</sup>

In 2014, a second genome cluster from *S. sp. MA37* was identified, and named *fdr*, since it is involved in the formation of 5-FDR.<sup>1</sup> This cluster is located separately from the *flA* cluster. It is similar to a section of the genome that encodes the chlorinase pathway in *Salinispora tropica*.<sup>3,1</sup>

The homology between the proteins found within the *fdr* cluster in *S. sp. MA37*, and their expected functions, along with the homology to *S. tropica* are displayed in **Fig. 3**.<sup>1</sup>



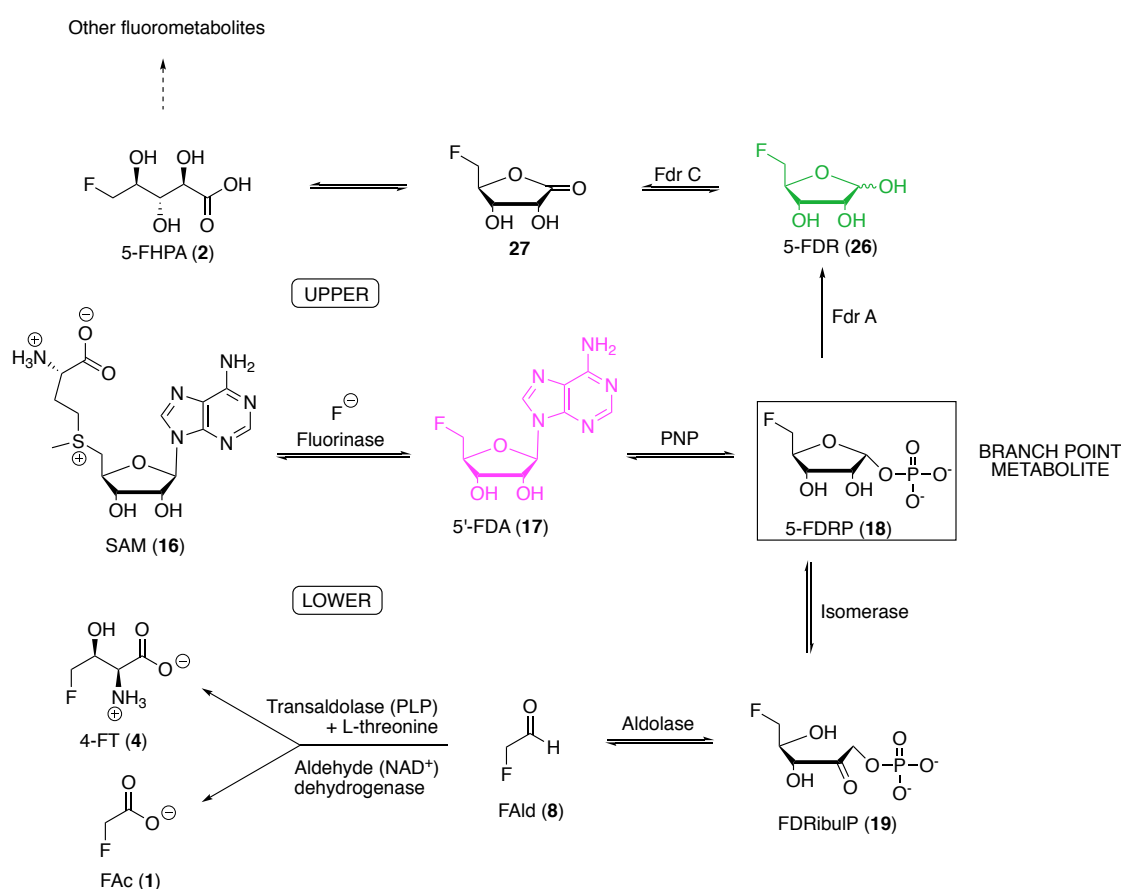
Protein	Residues	Deduced function	Homologue in <i>S. tropica</i>	% Identity
FdrG	309	ABC Transporter	-	-
FdrF	268	Ribulose-5-phosphate-4-epimerase	-	-
FdrA	267	Phosphoesterase	SalN	56 %
FdrB	259	Dihydroxy-acid dehydratase	SalH	68 %
FdrC	262	Short-chain dehydrogenase	SalM	69 %
FdrD	286	Cyclase	SalO	61 %
FdrE	213	LuxR family regulator	SalRII	38 %

**Fig. 3** *Fdr* cluster, and homology to *S. tropica*<sup>1</sup>

## 4.2 Aims and objectives

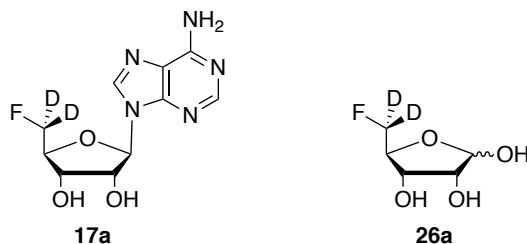
It became a focus to explore further the biosynthesis of the novel fluorometabolites of *Streptomyces* sp. MA37. Our approach was to incubate advanced biosynthetic intermediates with cell-free extracts, to track the pathway of previously unidentified fluorometabolites and gain a better knowledge of their origin.

For these purposes, it became an objective to synthesise candidate intermediates, for cell-free extract incubations. Metabolites immediately prior and after the putative branch point in the metabolic pathway (**Scheme 1**) were selected.



**Scheme 1.** Putative fluorometabolic pathway in *S. sp.* MA37. Target molecules for incubation studies are highlighted in pink and green

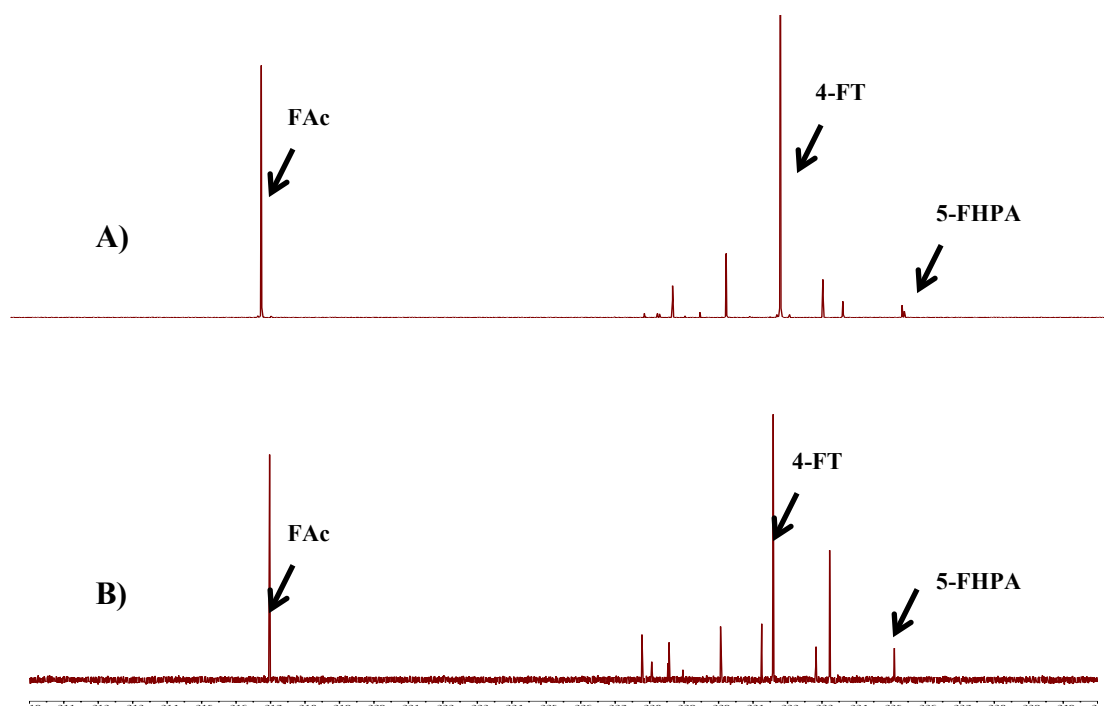
Accordingly, 5'-FDA (17) and 5-FDR (26) were synthesised, as indicated in **Section 3.3.2**. In order to further investigate the validity of the pathway, the syntheses of isotopically labelled  $[5',5'\text{-}^2\text{H}_2]$ -5'-fluoro-5'-deoxyadenosine **17a** and  $[5,5\text{-}^2\text{H}_2]$ -5-fluoro-5-deoxyribose **26a** also became an aim (**Fig. 4**).



**Fig. 4** Target isotopomers for feeding experiments

### 4.3 Fluorometabolite profile in *Streptomyces* sp. MA37

As a frame of reference, it became necessary to re-establish the production of the fluorometabolites from glycerol stocks of *S. sp.* MA37. Additionally, our collaborator Dr. H. Deng (University of Aberdeen), kindly supplied a fresh agar gel with plated bacteria. These strains were established in culture and both showed similar metabolite production when incubated in a fluoride-enriched medium. An example of the resultant  $^{19}\text{F}\{^1\text{H}\}$  NMR spectra of these extracts is shown in **Fig. 5**. Fluorometabolite production was good, and differed only a little from previous.<sup>1</sup>



**Fig. 5** A) Fluorometabolite production of a liquid culture of *S. sp.* MA37, when incubated in a fluoride-enriched medium, produced in this study. B) Original production of fluorometabolites in *S. sp.* MA37, as published<sup>1</sup>



This fluorometabolite profile will form the frame of reference for this study, where a number of cell-free extracts and whole-cell feeding experiments will be carried out to explore metabolite production.

#### **4.4 Cell-free experiments with *Streptomyces* sp. MA37**

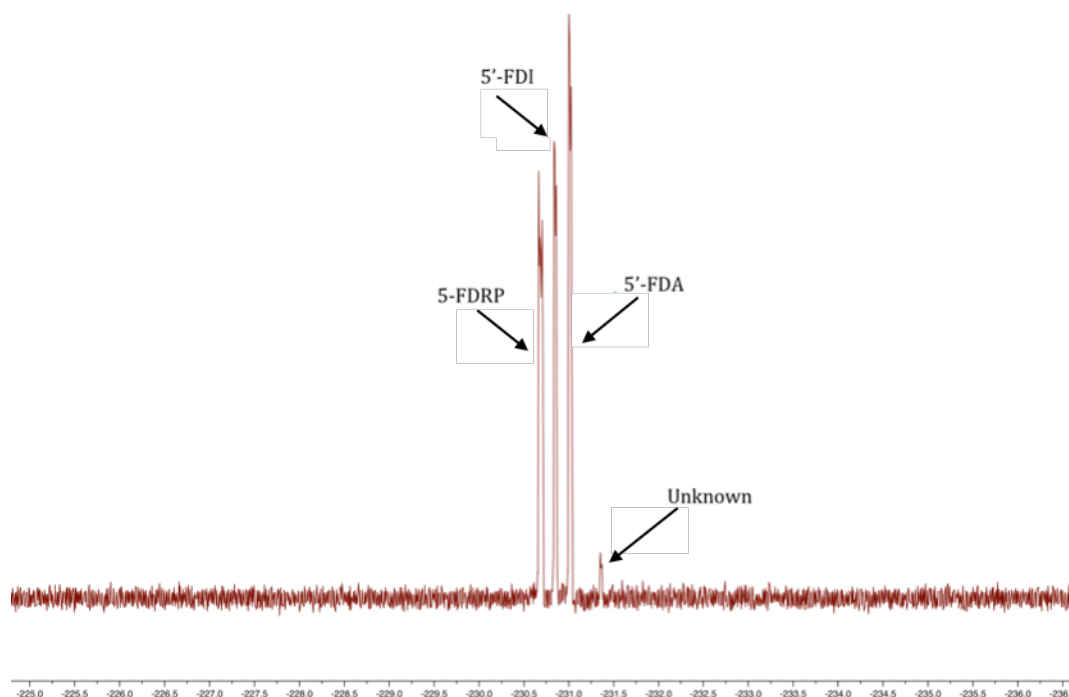
A ‘cell-free’ extract (CFE) refers to an experiment in which the cells have been ruptured, and the cell debris removed by centrifugation, but all the soluble elements inside the cell remain in the supernatant. This makes it possible to carry out incubation experiments with the released soluble, active enzymes.

##### **4.4.1 Exploratory incubations of SAM with inorganic fluoride, and 5'-FDA. Assignment of signals**

A cell-free extract was prepared following the procedure described in **Section 5.7.2 (Chapter 5)**. SAM (**16**) and 5'-FDA (**17**) were stored in 10 mM stock solutions and KF in a 50 mM solution, at -80 °C. Reagent solutions were thawed in ice before use. A first set of experiments was carried out incubating with added SAM (in the presence of fluoride ion) to assess fluorinase activity in the CFE. Also, 5'-FDA (**17**) was explored as a substrate to assess if the downstream metabolic enzymes remained active in the CFE.

The SAM (**16**) and fluoride experiments were run for 16–48 h, but they did not show any fluorine peaks by  $^{19}\text{F}\{^1\text{H}\}$  NMR spectroscopy, apart from that corresponding to fluoride ion. This lack of fluorometabolite production suggested that the fluorinase in this particular CFE was not active.

However, incubation with 5'-FDA (**17**) did generate a metabolite profile, as identified by  $^{19}\text{F}\{^1\text{H}\}$  NMR spectroscopy. A synthetic sample of 5'-FDA was added into the NMR tube. This confirmed that the signal at -231.0 ppm was indeed 5'-FDA. Given the proximity of these peaks within the spectrum – all appearing within a 0.5 ppm range – it is difficult to assign the peaks accurately without spiking with reference compounds. The  $^{19}\text{F}$  NMR peaks at -230.7, -230.9, and -231.0 ppm could be assigned to 5-FDRP (**18**), 5'-FDI (**153**) and 5'-FDA (**17**), respectively.<sup>4</sup> The peak at -231.4 ppm is an unknown metabolite.

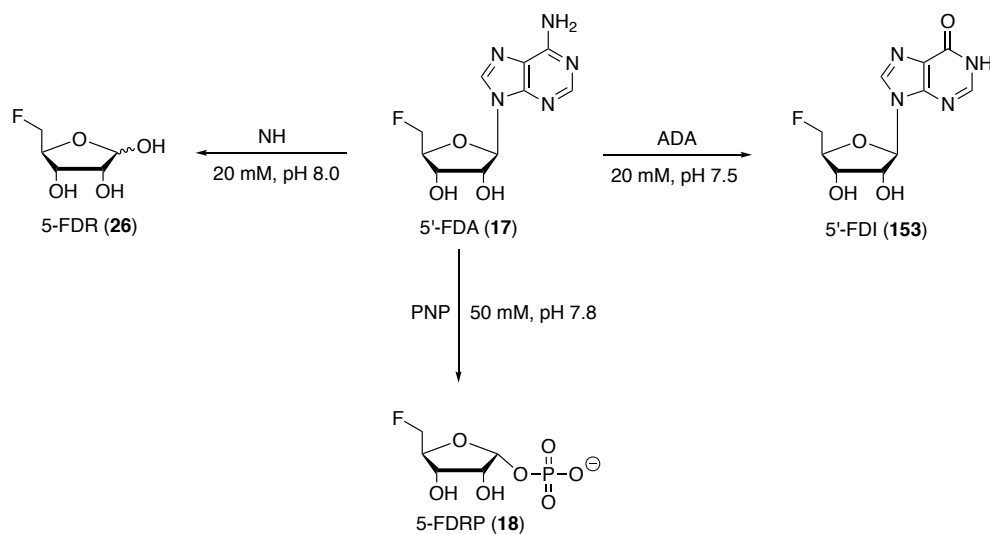


**Fig. 6**  $^{19}\text{F}\{^1\text{H}\}$  NMR spectrum of the 5'-FDA incubation in CFE in phosphate buffer, note the false multiplicity of the peaks, caused by faulty shimming in the NMR spectroscopy

Given the clustering of the fluorine peaks within the spectra, an internal  $^{19}\text{F}$ -NMR reference compound was added to all of the cell free extracts to improve confidence in signal assignments. 1,3-Difluoro-2-propanol was selected due to its high solubility in water, and its proximity in chemical shift (-232.8 ppm). The internal reference revealed that the peaks did not always appear at the same chemical shift from different experiments, most probably due to differences in the protein concentration or pH of the different samples.

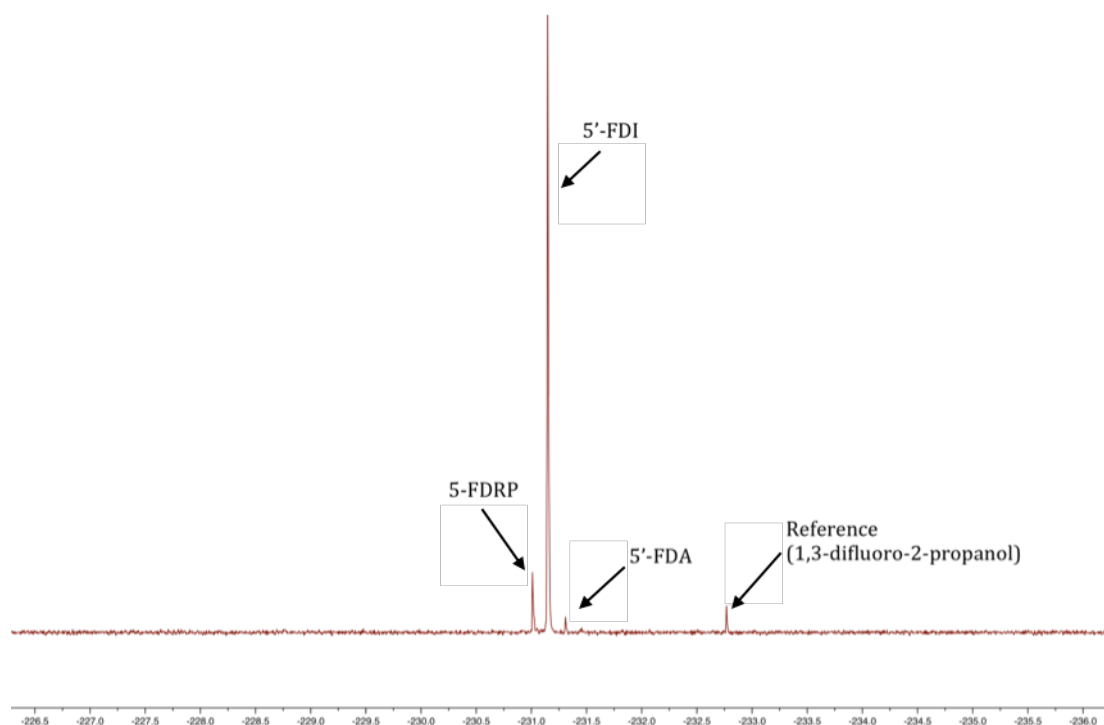
Since, no synthetic reference samples of 5'-FDI (**153**) and 5-FDRP (**18**) were available, it proved necessary to prepare samples of 5'-FDI and 5-FDRP by enzymatic biotransformations. In addition, the same nucleoside hydrolase (NH) enzyme used in the Chapter 3 was employed to generate a sample of 5-FDR (**26**).

Three enzymatic transformations were carried out with 5'-FDA (**17**), as shown in **Scheme 2**.<sup>5</sup> The adenosine deaminase (ADA) and hydrolase (NH) reactions worked as expected, affording 5'-FDI (**153**) and 5-FDR (**26**), respectively.



**Scheme 2.** Enzymatic transformations performed on 5'-FDA (17)

The purine nucleoside phosphorylase (PNP) reaction led to three fluorinated products (in the  $^{19}\text{F}\{^1\text{H}\}$  NMR spectrum with chemical shifts at -231.0, -231.1, and -231.3 ppm, **Fig. 7**). 5'-FDI **153** (-231.1 ppm) and 5'-FDA **17** (-231.3 ppm) were established as the major products, and it was deduced by comparison to previous reports,<sup>7</sup> that the minor signal at -231.0 ppm belonged to 5-FDRP **18** (**Fig. 7**).

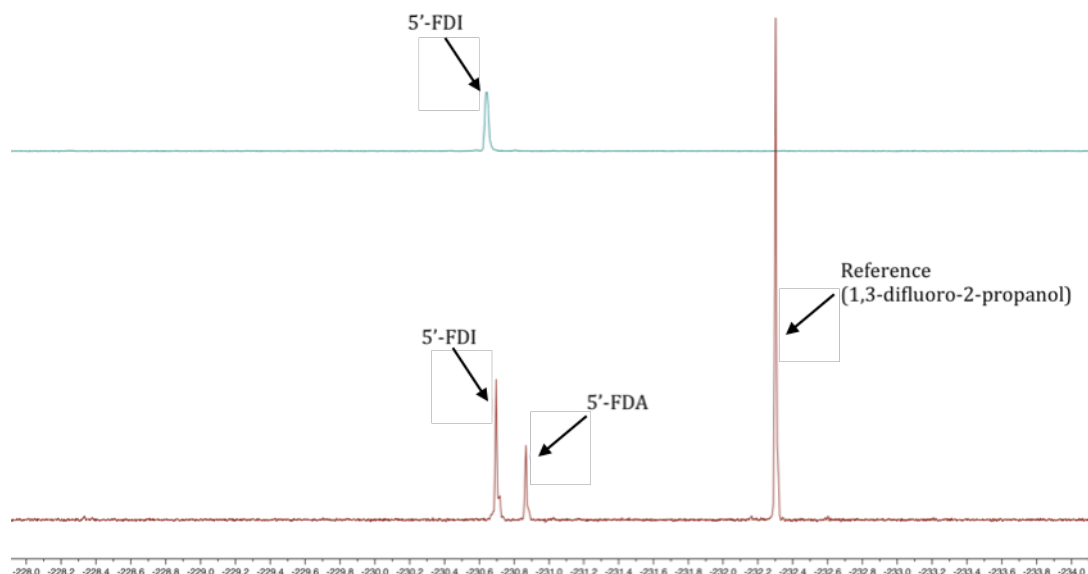


**Fig. 7**  $^{19}\text{F}\{^1\text{H}\}$  NMR spectrum for the PNP reaction with 5'-FDA, after 5'-FDA spiking

The products obtained in these enzymatic reactions were used as reference compounds for spiking the CFEs. With these reference compounds in hand, it was established that the three major peaks obtained when incubating 5'-FDA in the CFE were (from left to right) 5-FDRP, 5'-FDI and 5'-FDA, as illustrated in **Fig. 7**.

The cell-free experiments carried out so far indicated that the enzyme activity beyond the branch point metabolite 5-FDRP seemed not to be active in the incubations. Furthermore, there did not appear to be any significant fluorinase activity. In order to confirm this, these experiment were repeated. A cell pellet from a new culture was prepared, and the CFE experiments with both SAM/KF and 5'-FDA were repeated. The results were similar to those carried out with the first pellet. The time (day 8) at which the cell pellet was harvested was also investigated. Two fresh pellets were prepared on days 14 and 20, respectively. The outcome after incubating SAM/KF and 5'-FDA with the resultant CFEs did not change.

A different buffer for the CFE preparation was explored. The experiments were run this time in Tris (20 mM, pH 7.5), rather than phosphate buffer. Incubation of fluoride ion and SAM did not progress, but incubation with 5'-FDA this time showed only one product by  $^{19}\text{F}$  NMR spectroscopy, which could be assigned to 5'-FDI (**153**). **Fig. 8** shows  $^{19}\text{F}\{^1\text{H}\}$  NMR spectra for the experiment (upper spectrum) and the same experiment after spiking (lower spectrum) with the reference compound (1,3-difluoro-2-propanol, -232.8 ppm) and 5'-FDA (-231.2 ppm).



**Fig. 8**  $^{19}\text{F}\{^1\text{H}\}$  NMR spectra of the 5'-FDA CFE run in Tris buffer (upper), and the same sample after spiking with 5'-FDA and the reference (lower)

In order to probe 5'-FDI (**153**) as an intermediate metabolite, a cell-free incubation was prepared, adding 5'-FDI. However, **153** did not progress to any other fluorometabolites, either in phosphate or Tris buffers, consistent with its presence as a dead end 'shunt' metabolite.

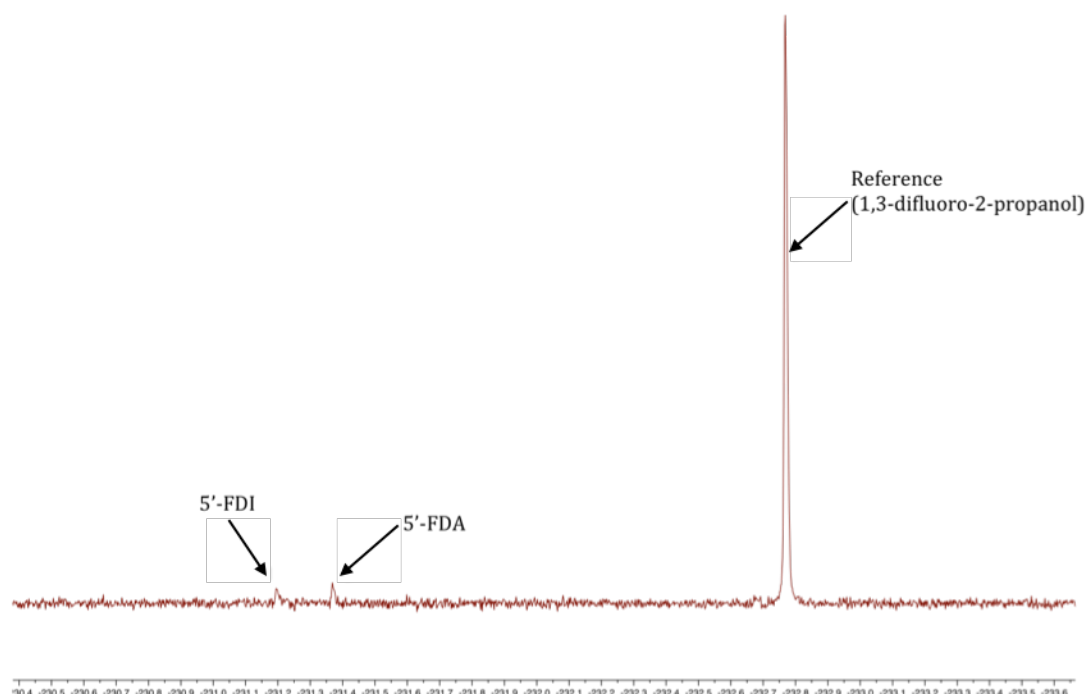
In all of the cell-free experiments, the adenosine deaminase (ADA) enzyme, which converts 5'-FDA to 5'-FDI, proved to be more active than the PNP. Inosine was added into the CFE, in an attempt to inhibit the deaminase, however, the formation of 5'-FDI still dominated, and no new fluorometabolite peaks were obtained.

The lack of 5-FDRP (and the increased deaminase activity of PNP) in Tris buffer was not unexpected, since phosphate ions are required for PNP activity.

#### 4.4.2 Effect of the addition of fluorinase and influence of 5'-FDA concentration

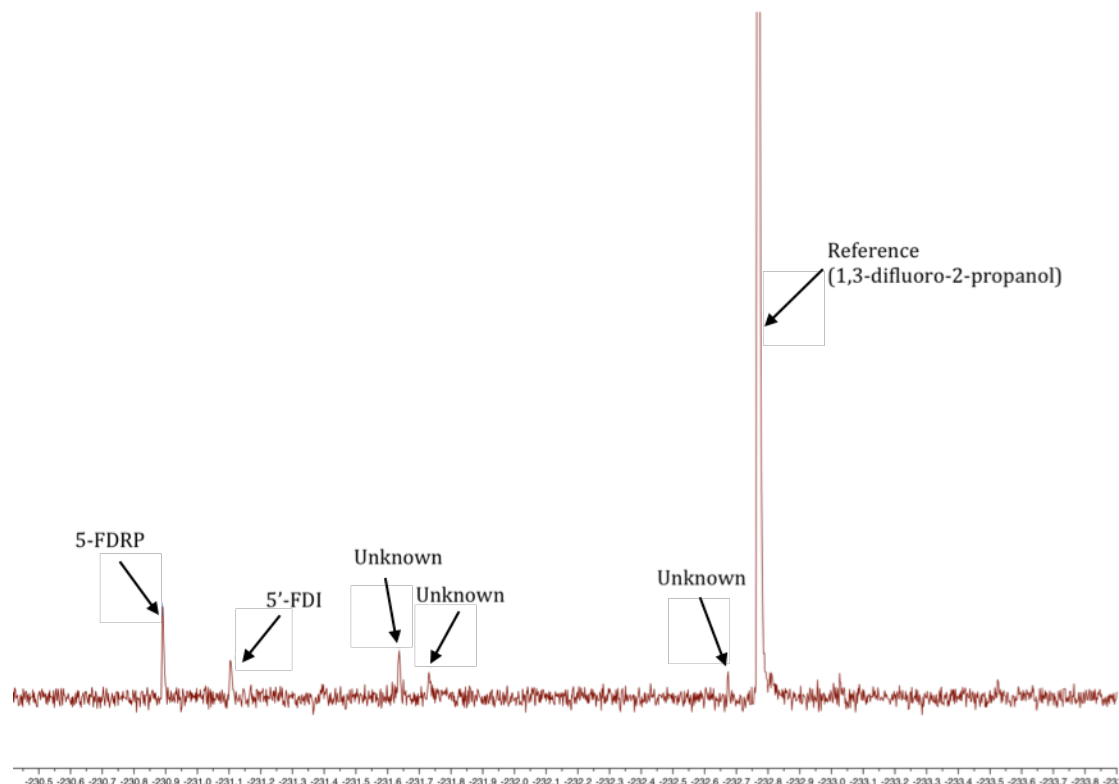
Since fluorinase in the cell-free incubations was inactive, CFE experiments were explored with added recombinant fluorinase. With such a combination, incubation of SAM and fluoride led to the formation, for the first time, of 5'-FDA (**17**) and 5'-FDI

(153), **Fig. 9**. This indicated again that there was still significant adenosine deaminase activity in the CFE.



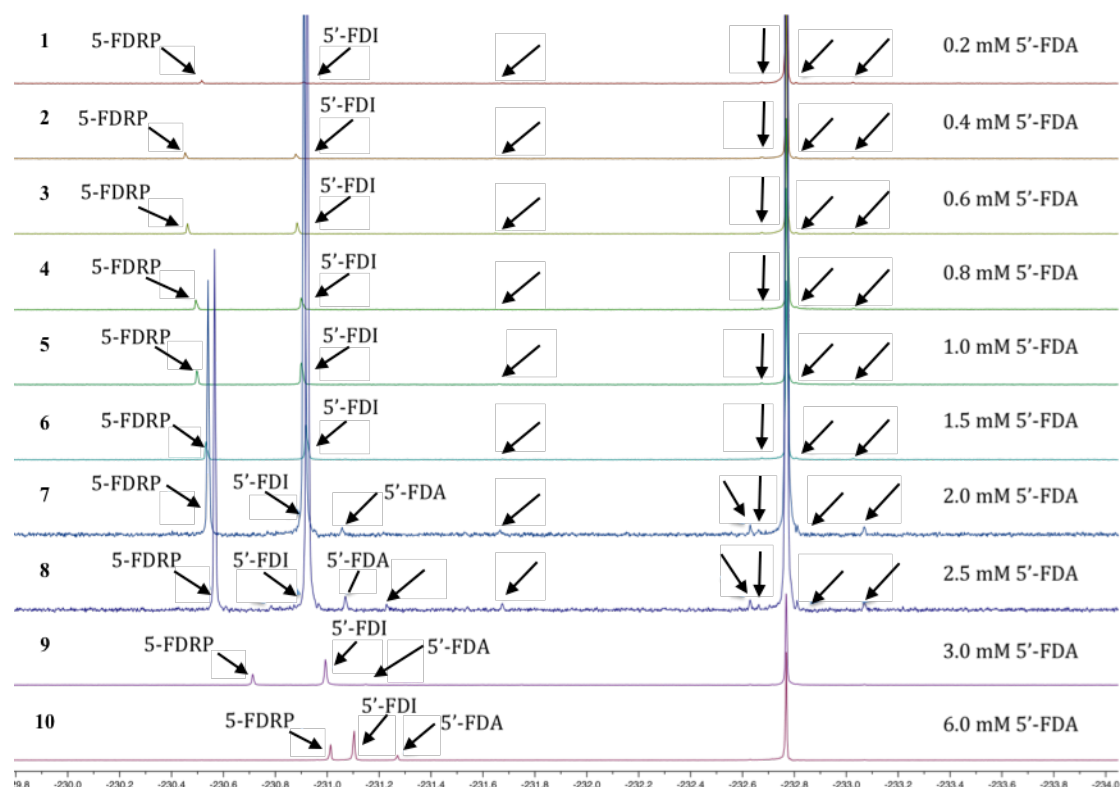
**Fig. 9**  $^{19}\text{F}\{^1\text{H}\}$  NMR spectrum of SAM/KF CFE, with a recombinant fluorinase

The final concentration of 5'-FDA (**17**) in previous CFEs was 3 mM, which may lead to feedback inhibition. Thus, a new CFE was carried out with a final concentration of just 0.5 mM of 5'-FDA. This was more successful and new peaks appeared at -230.9, -231.1, -231.6, -231.7, -232.7, -233.0 and -233.5 ppm in the  $^{19}\text{F}\{^1\text{H}\}$  NMR spectrum (**Fig. 10**). This was the first time that a wide range of fluorometabolites accumulated in a cell-free extract with 5'-FDA, and suggests that higher concentrations of 5'-FDA (**17**) shut down the biotransformations. The identity of many of the peaks in **Fig. 10** remains unknown.



**Fig. 10**  $^{19}\text{F}\{^1\text{H}\}$  NMR spectrum of the 5'-FDA CFE at 0.5 mM

In order to obtain more information regarding the effect of 5'-FDA (**17**) concentration in the CFEs, a study was carried out incubating 5'-FDA across a range of concentrations; 0.2, 0.4, 0.6, 0.8, 1.0, 1.5, 2.0, 2.5, 3 and 6 mM were selected. As shown in **Fig. 11**, concentrations below 3.0 mM seem to be more active, generating a wider range of fluorometabolites, while the assays performed at 3.0 and 6.0 mM show a clear predominance of 5-FDRP, 5'-FDI and remaining 5'-FDA.

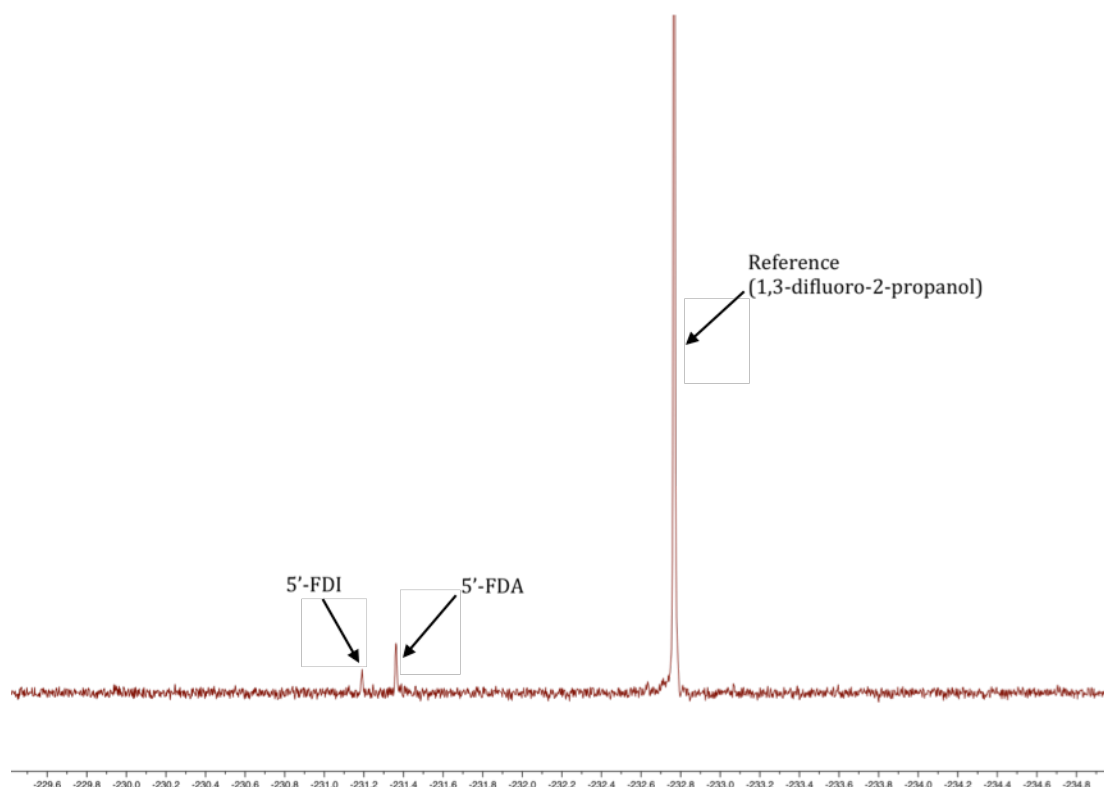


**Fig. 11**  $^{19}\text{F}\{^1\text{H}\}$  NMR spectra of the 5'-FDA concentration (0.2 – 6.0 mM) assays

**Fig. 11** shows an important shift in the peaks, especially in the experiments with 3.0 and 6.0 mM 5'-FDA concentration. The pH of the solution in the NMR tubes remained the same, so these shifts are mostly due to a difference in the protein concentration. In order to make sure that the peaks correspond to the same compounds, NMR tubes of experiment 4 (0.8 mM) and experiment 10 (6.0 mM) were mixed. The outcome showed just one set of peaks, indicating that the products were the same in both experiments.

In a similar line of investigation, the effects of the concentration of SAM and fluoride were investigated. In these experiments, a concentration ratio of 1:5 mM of SAM to an excess of fluoride was used (previously the ratio SAM : KF was 1:1). For the first time, fluorometabolites were observed without addition of a recombinant fluorinase. This also led to the formation of 5'-FDA and 5'-FDI, as shown in **Fig. 12**.





**Fig. 12**  $^{19}\text{F}\{^1\text{H}\}$  NMR spectrum of the SAM and KF CFE, keeping the ratio 1:5 mM

#### 4.4.3 Addition of cofactors, supplements and enzymes into the CFE incubations

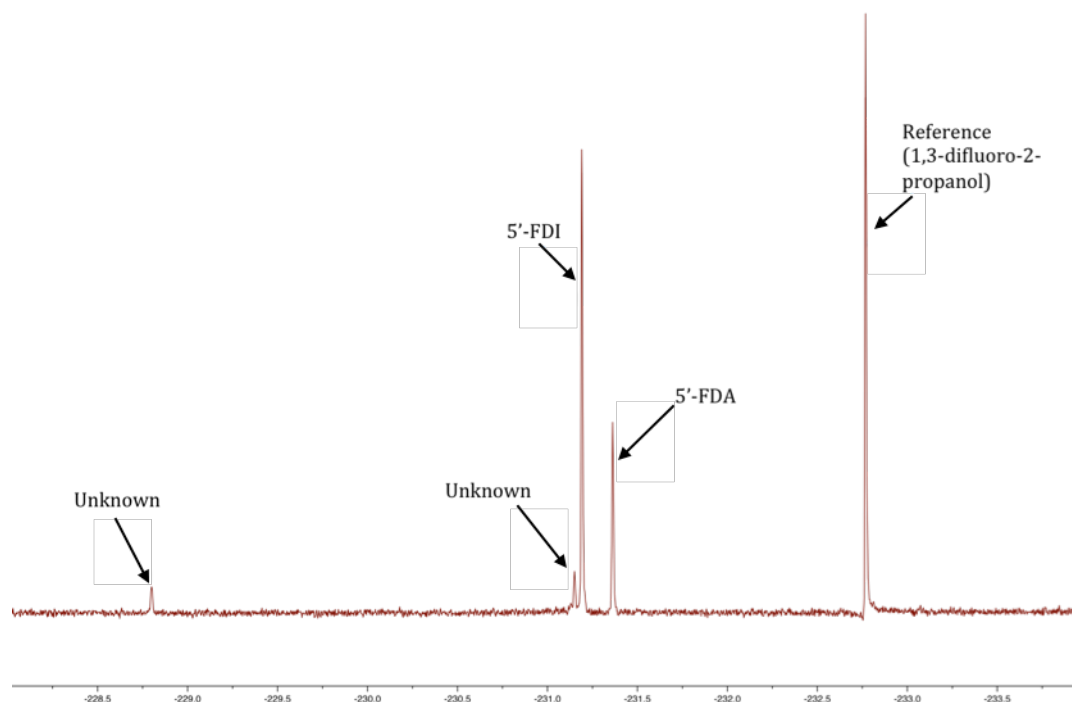
Given that some of the structures and their metabolic pathways are still unknown, there is the possibility that the CFEs were not fully competent because the enzymes were lacking required cofactors.

A range of experiments with different cofactors (**Table 2**), buffers and feedstocks were trialled. A summary of the experimental conditions used, and corresponding outcomes are given in **Table 2**. All of the CFE experiments were carried out with a pellet harvested on day 8 of the culture fermentation, and the CFE incubations were performed for 20 h at 37 °C. For the incubations involving SAM and fluoride, a recombinant fluorinase was added (30  $\mu\text{L}$  of a 2.2 mg/mL solution), in an attempt to stimulate the production of fluorometabolites.

**Table 2.** Addition of different enzymes, cofactors and supplements, in both Tris·HCl and phosphate buffers

Buffer	Enzyme	Cofactor	Supplement	Substrate	Results
Tris · HCl	Fluorinase	PLP	-	SAM/KF	-
Phosphate	Fluorinase	PLP	-	SAM/KF	-
Tris · HCl	-	PLP	-	FDA	5'-FDI, 5'-FDA,
Phosphate	-	PLP	-	FDA	5'-FDI, 5-FDRP, 5'-FDA
Tris · HCl	Fluorinase	NAD <sup>+</sup>	-	SAM/KF	-
Phosphate	Fluorinase	NAD <sup>+</sup>	-	SAM/KF	-
Tris · HCl	-	NAD <sup>+</sup>	-	FDA	5'-FDI, 2 unknown
Phosphate	-	NAD <sup>+</sup>	-	FDA	5'-FDI, 5-FDRP, 5'-FDA
Tris · HCl	Fluorinase	ATP	Mg <sup>2+</sup>	SAM/KF	-
Phosphate	Fluorinase	ATP	Mg <sup>2+</sup>	SAM/KF	-
Tris · HCl	-	ATP	Mg <sup>2+</sup>	FDA	5'-FDI, 5'-FDA
Phosphate	-	ATP	Mg <sup>2+</sup>	FDA	5'-FDI, 5-FDRP
Tris · HCl	Fluorinase	FAD	-	SAM/KF	-
Phosphate	Fluorinase	FAD	-	SAM/KF	-
Tris · HCl	-	FAD	-	FDA	5'-FDI, 2 unknown peaks
Phosphate	-	FAD	-	FDA	5'-FDI, 5-FDRP
Tris · HCl	Fluorinase	NAD <sup>+</sup> , FAD, ATP, PLP	Mg <sup>2+</sup>	SAM/KF	-
Phosphate	Fluorinase	NAD <sup>+</sup> , FAD, ATP, PLP	Mg <sup>2+</sup>	SAM/KF	-
Tris · HCl	-	NAD <sup>+</sup> , FAD, ATP, PLP	Mg <sup>2+</sup>	FDA	5'-FDI, 2 unknown peaks
Phosphate	-	NAD <sup>+</sup> , FAD, ATP, PLP	Mg <sup>2+</sup>	FDA	5'-FDI, 5-FDRP

In these assays, a new peak emerged at -228.8 ppm (consistent with the chemical shift of  $\beta$ -5-FDR) when 5'-FDA was incubated. There is no 5-FDRP present, but there is a second peak appearing at -231.3 ppm (**Fig. 13**). Since all of the experiments with cofactors and 5'-FDA led to the formation of the new peak at -228.8 ppm, it is probable that this peak appears as a feature of the particular culture used.

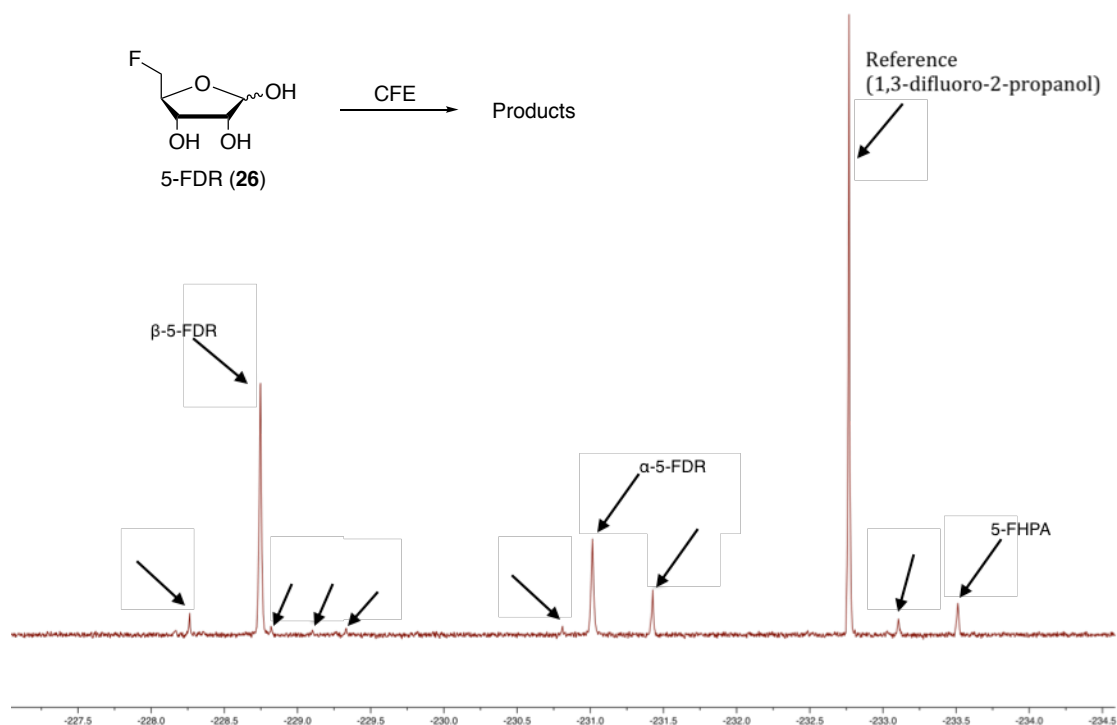


**Fig. 13**  $^{19}\text{F}\{^1\text{H}\}$  NMR spectrum of a standard sample taken from one of the CFE incubations of 5'-FDA with PLP in Tris buffer

Despite the incubations with 5'-FDA leading to the production of fluorometabolites, none of the assays carried out with SAM and KF generated any; and none of the cofactors added was responsible for the generation of new fluorometabolites.

#### 4.4.4 Incubation of 5-FDR. Reproduction of published experiments.

Incubation of 5-fluoro-5-deoxyribose (5-FDR, **26**) with a newly prepared cell-free extract gave, as expected, a range of fluorometabolites as observed previously.<sup>1</sup> No fluoroacetate or 4-fluorothreonine were observed, a result entirely consistent with a branching pathway. The results obtained in this incubation are shown in **Fig. 14**.

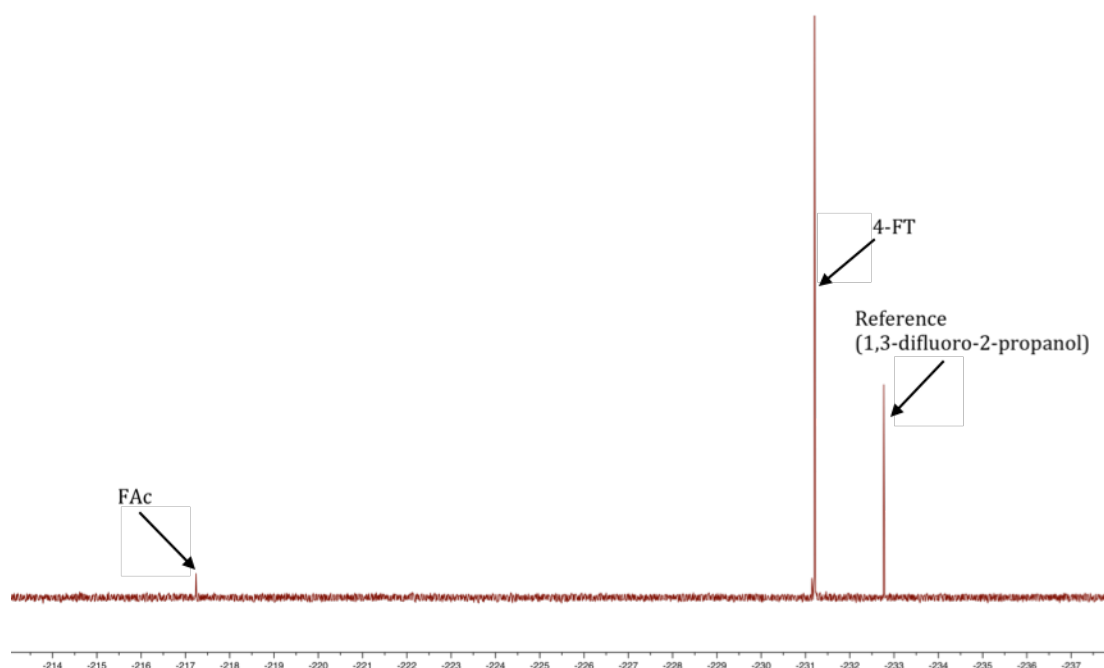


**Fig. 14** <sup>19</sup>F{<sup>1</sup>H} NMR spectrum of the 5-FDR CFE incubation in phosphate buffer

#### 4.5 Incubation in cell-free experiments from *Streptomyces cattleya*

A cell-free experiment was carried out in order to check that the protocols followed in this study for CFE preparation were valid. This involved a repetition of the CFE experiments previously reported for the historical fluoroacetate and 4-fluorothreonine producing organism *S. cattleya*,<sup>6</sup> and involved addition of 5'-FDA (**17**).

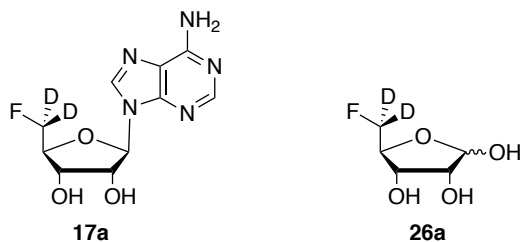
This experiment led to the efficient formation of 4-FT **4** and FAc **1**, as previously observed and gave confidence that the protocols were correctly carried out. In overview, these experiments suggest that *Streptomyces* sp. MA37 generates a much more sensitive CFE than that from *S. cattleya*.



**Fig. 15**  $^{19}\text{F}\{^1\text{H}\}$  NMR spectrum of the 5'-FDA CFE incubation in *S. cattleya*

#### 4.6 Synthesis of the deuterium labelled intermediates for feeding experiments

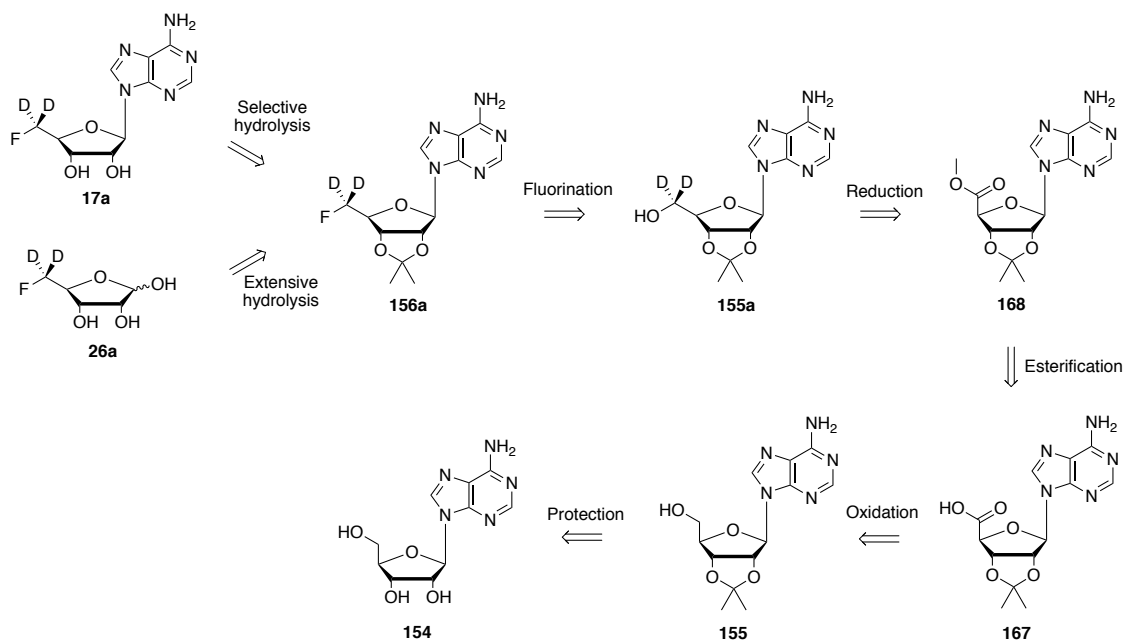
As well as feeding experiments with 5'-FDA (**17**) and 5-FDR (**26**), prepared in Chapter 3, we envisaged the preparation of deuterium-labelled intermediates, as such compounds can facilitate tracking the origin of the fluorometabolites. Deuterium incorporations are readily identified by a heavy atom induced shift in the fluorine signals of the  $^{19}\text{F}$ -NMR spectrum, when compared to their non-deuterated analogues.<sup>7</sup> Therefore, a comparison between the spectra produced in the labelled and unlabelled experiments, should reveal isotopically enriched metabolites, which derived from the labelled precursor. Hence,  $[5',5'\text{-}^2\text{H}_2]\text{-5'-fluoro-5'-deoxyadenosine}$  **17a** and  $[5,5\text{-}^2\text{H}_2]\text{-5-fluoro-5-deoxyribose}$  **26a** became synthetic targets.



These compounds were selected as they are intermediate metabolites, which occur just before and after the proposed metabolic branch point. This branch occurs at 5-fluoro-5-deoxyribose phosphate **18**, as highlighted in **Scheme 1** (page 171). The synthesis of [5',5'- $^2\text{H}_2$ ]-5'-fluoro-5'-deoxyadenosine **17a** should feed into both metabolic branches, whereas [5,5- $^2\text{H}_2$ ]-5-fluoro-5-deoxyribose **26a** should only label compounds derived from the upper branch of the pathway, if the general hypothesis is correct.

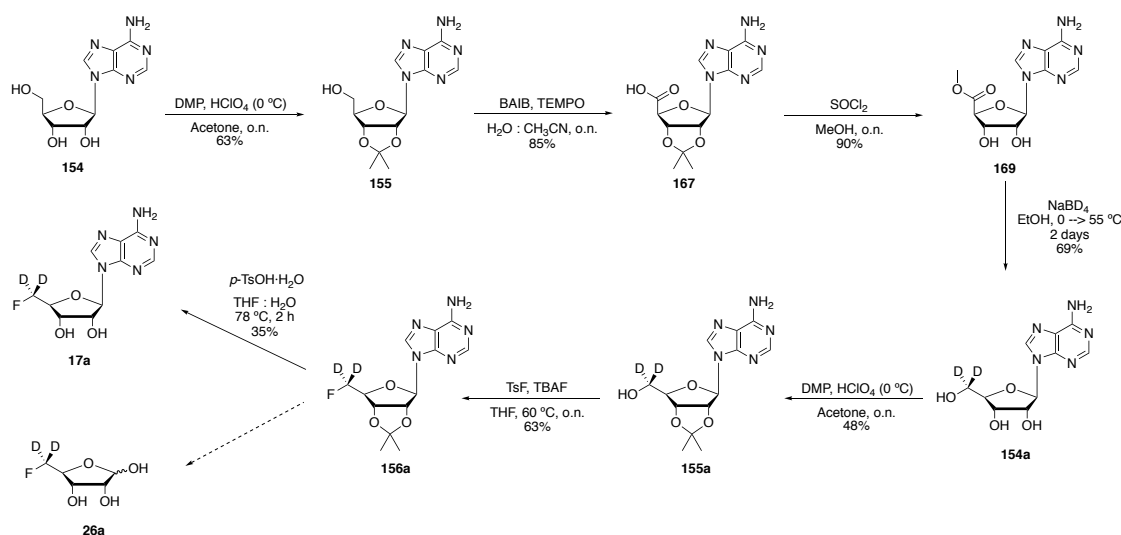
#### 4.6.1 Synthesis of [5',5'- $^2\text{H}_2$ ]-5'-fluoro-5'-deoxyadenosine (**17a**) and [5,5- $^2\text{H}_2$ ]-5-fluoro-5-deoxyribose (**26a**)

Given the very similar location of the isotope in the fluoromethyl groups of these two compounds, a common route was designed with modification only of the last step. The retrosynthesis is shown in **Scheme 3**.



**Scheme 3.** Retrosynthesis of the isotopically labelled precursors

The proposed synthesis began with the protection of adenosine **154** as an acetonide, followed by oxidation to a carboxylic acid and then subsequent esterification. The deuterium isotope will be introduced by reduction of the ester moiety in the presence of a deuteride source. Fluorination and deprotection constitute the final steps. Varying the conditions of this last step should allow for the formation of either **17a** or **26a**.

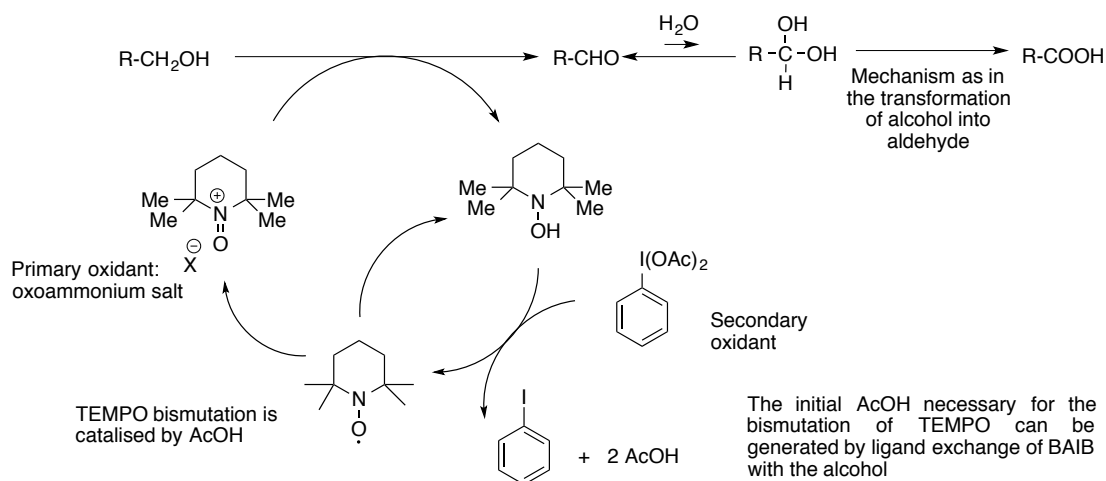


**Scheme 4.** Conditions and reagents used in the synthesis of [5',5'- $^2\text{H}_2$ ]-5'-fluoro-5'-deoxyadenosine **17a** and [5,5- $^2\text{H}_2$ ]-5-fluoro-5-deoxyribose **26a**

Protection of adenosine **154** was carried out in the presence of 2,2-dimethoxypropane (DMP), perchloric acid and acetone. The addition of both DMP and perchloric acid was performed at 0 °C, and the latter was added dropwise. After purification, this gave acetonide **155** in 63% yield.

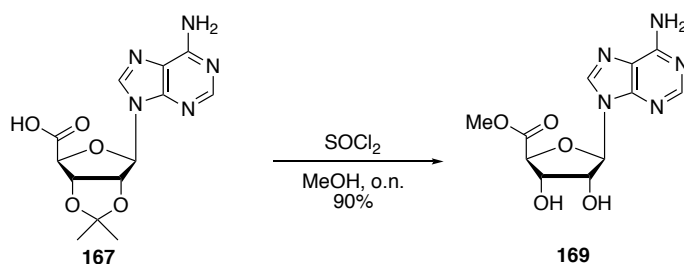
The oxidation of **155** into carboxylic acid **167** with BAIB and TEMPO in a 1:1 mixture of MeCN/water proved straightforward and **167** was precipitated from the reaction mixture in an 85% yield.

This oxidation reaction was previously described for substrate **155** by Epp and Widlanski in 1999,<sup>8</sup> and the mechanism, which proceeds by a radical process, is described by Tojo and Fernández.<sup>9</sup>



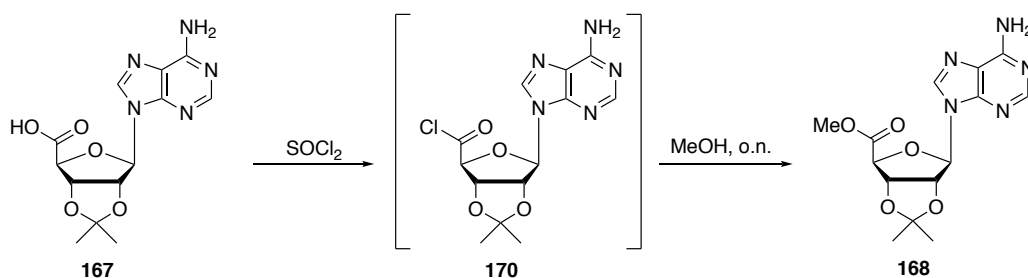
**Scheme 5.** Mechanism for the oxidation, as described by Tojo and Fernández<sup>12</sup>

The esterification reaction from **167** to **168** using  $\text{SOCl}_2$  in MeOH was not straightforward. The acetonide protecting group was hydrolysed during the procedure, due to the acidic conditions, generating adenosine-5'-carboxylic acid methyl ester **169** in a 90% yield.



**Scheme 6.** Formation of denosine-5'-carboxylic acid methyl ester **169**

Esterification conditions for this substrate **167**, while minimising hydrolysis of the acetonide group, were described by Brondyk in 1976.<sup>10</sup> The conditions used were the same, but the reaction was carried out in a one-pot, two-step procedure, as shown **Scheme 7**.

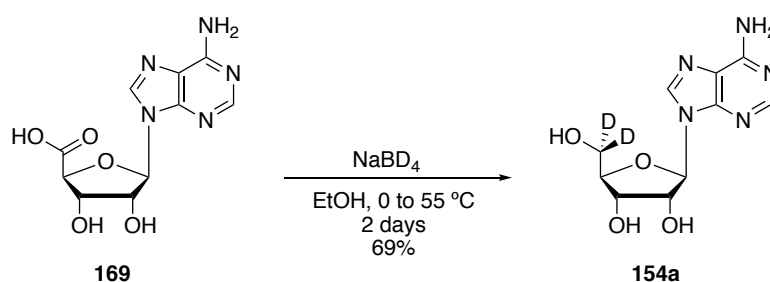


**Scheme 7.** Brondyk conditions for the esterification of **168**<sup>13</sup>



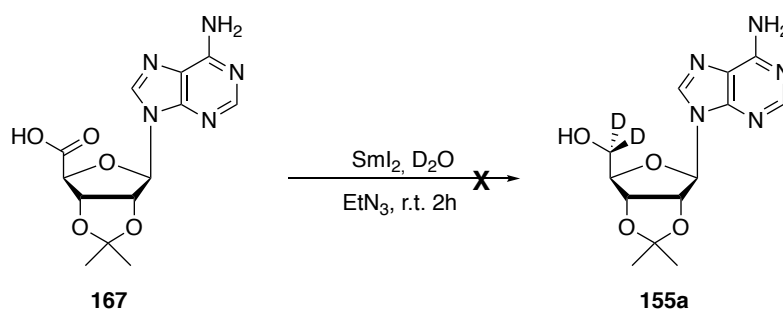
These new conditions worked relatively well on a small scale (100 mg of starting material), giving **168** in a 46% yield. However, when the reaction was scaled up, product isolation was hampered by a significant level of deprotected ester **169**. It proved more efficient to progress with the deprotected adenosine **169**.

The next reaction of the sequence involved the reduction of ester **169** in the presence of a deuteride source ( $\text{NaBD}_4$ ). This afforded di-deuterated adenosine **154a** in 69% yield.



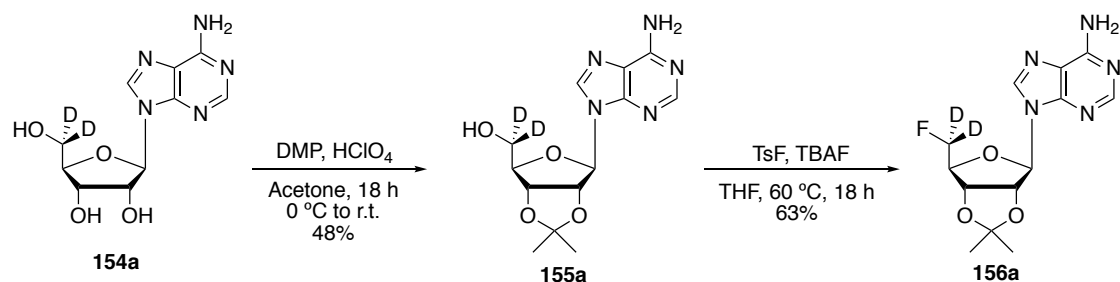
**Scheme 8.** Reduction of the ester **169** with  $\text{NaBD}_4$

Alternatively, in an attempt to reduce the synthesis by two steps, and therefore improve the overall yield, conditions were explored to transform carboxylic acid **167** into the di-deuterated and protected adenosine **155a**, as described by Szostak and Procter in 2014.<sup>11</sup> However, the reaction did not work, most likely due to the lack of solubility of **167** in  $\text{D}_2\text{O}$ , which was both the solvent and the deuterium source.



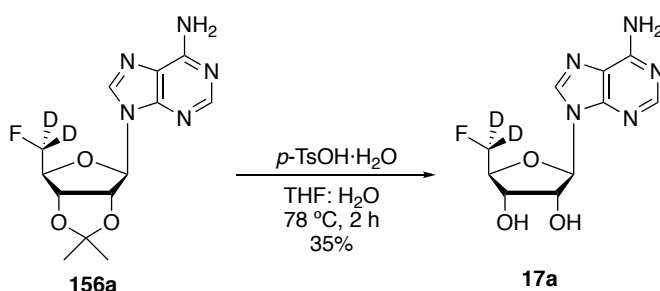
**Scheme 9.** Attempted  $\text{SmI}_2$  reduction of carboxylic acid **155a**

The original synthetic plan required the re-protection of  $[5',5'\text{-}^2\text{H}_2]$ -adenosine **154a**, a reaction that was carried out following the previously described procedure. In the event, acetone **155a** was obtained in 48% yield, and then fluorination with TBAF and TsF in THF at 60 °C was performed. The reaction proved straightforward and led to **156a** in a 63% yield.



**Scheme 10.** Re-protection and fluorination reactions

Finally, deprotection was required. In the case of [5',5'- $^2\text{H}_2$ ]-5'-FDA **17a**, the procedure also proved straightforward, and product **17a** was obtained in a 35% yield.



**Scheme 11.** Deprotection of the acetonide to give [5'5'- $^2\text{H}_2$ ]-17 (**17a**)

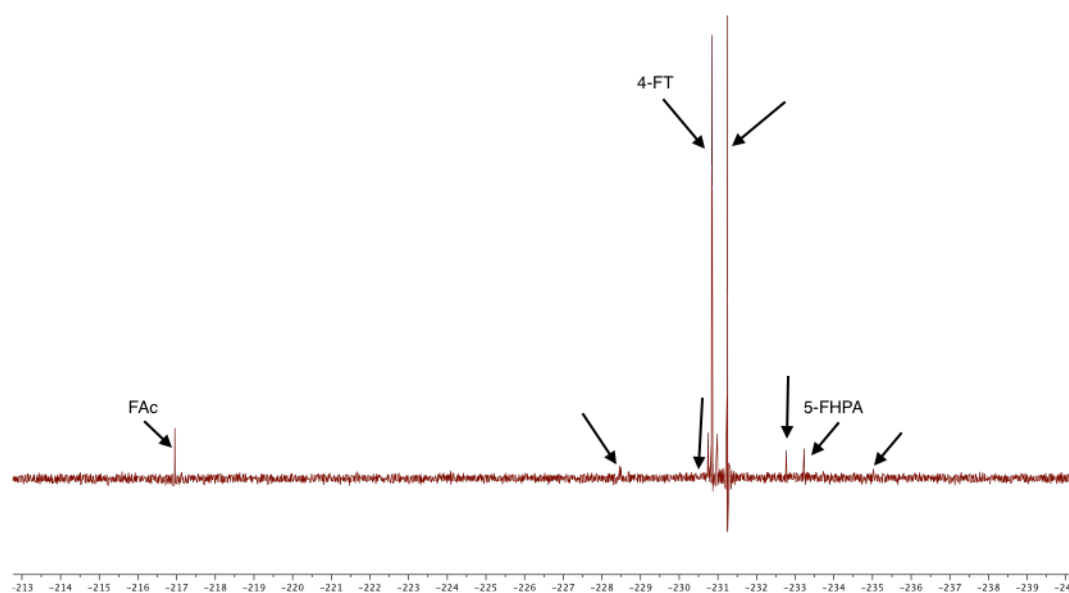
Deprotection to give [5,5- $^2\text{H}_2$ ]-FDR **26a** was monitored by  $^{19}\text{F}\{^1\text{H}\}$  NMR spectroscopy, and the formation of the anomers of [5,5- $^2\text{H}_2$ ]-FDR (**26a**) was observed after 2 h in 4M HCl. Purification of this highly water soluble product was challenging and it proved difficult to isolate a useful sample.

## 4.7 Feeding experiments in whole cell cultures of *S. sp.* MA37

### 4.7.1 Feeding of 5'-FDA (**17**)

In the first place, unlabelled 5'-FDA (**17**) was added into a whole cell culture as the sole source of fluorine. Two different experiments were carried out, to investigate whether the time of addition would influence the incorporation or production. Therefore, in one of the experiments 5'-FDA was added after inoculating the bacteria on day 0 of the cycle; whereas in the second experiment it was added at day 8, when fluorometabolite production starts, in a standard culture with inorganic fluoride in the medium. The experiments were then incubated for 20 days, under standard conditions.

Whilst the incubation inoculated at day 8 shows the formation of 4-fluorothreonine (**4**) and fluoroacetate (**1**) – along with residual 5'-FDA –; the incubation that had 5'-FDA (**17**) added from the start of the experiment produced a wider range of fluorometabolites (**Fig. 16**).



**Fig. 16** Feeding experiment in *S. sp.* MA37, after 18 days (inoculation of 5'-FDA on day 0)

As shown in **Fig. 16**, the production of fluorometabolites is similar to that observed during the profiling of *S. sp.* MA37, suggesting that whole cells are needed in order to generate the full fluorometabolite profile.

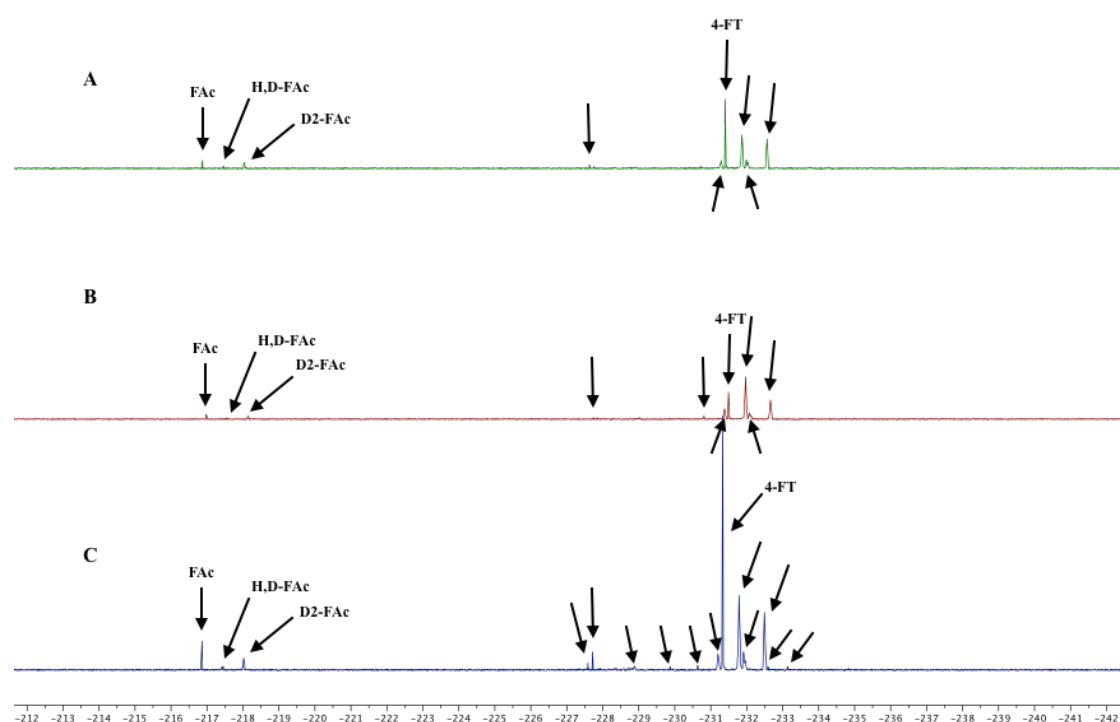
A control experiment was also run to confirm that fluorometabolite production is derived from the 5'-FDA present in the culture. For this, a culture of *S. sp.* MA37 was incubated for 25 days, under identical conditions, but without an external source of fluoride.  $^{19}\text{F}$  NMR spectra is clear and shows no fluorinated peaks.

#### 4.7.2 Feeding of [5',5'- $^2\text{H}_2$ ]-5'-FDA (**17a**)

After the results with unlabelled 5'-FDA **17**, incubations with deuterium-labelled [5',5'- $^2\text{H}_2$ ]-5'-FDA **17a** were explored.

For this, three experiments were conducted. In the first experiment  $[5',5'\text{-}^2\text{H}_2]\text{-5'-FDA}$  (**17a**) was added shortly after the inoculation of the media with *S. sp.* MA37, on day 0, and this was the only source of fluorine. In the second experiment,  $[5',5'\text{-}^2\text{H}_2]\text{-5'-FDA}$  (**17a**) was added sequentially on days 4, 5 and 6; again as the sole source of fluorine. The last experiment reproduced the latter, with sequential addition of  $[5',5'\text{-}^2\text{H}_2]\text{-5'-FDA}$  (**17a**) on days 4, 5 and 6, but it also contained KF (0.2 mM) in the medium.

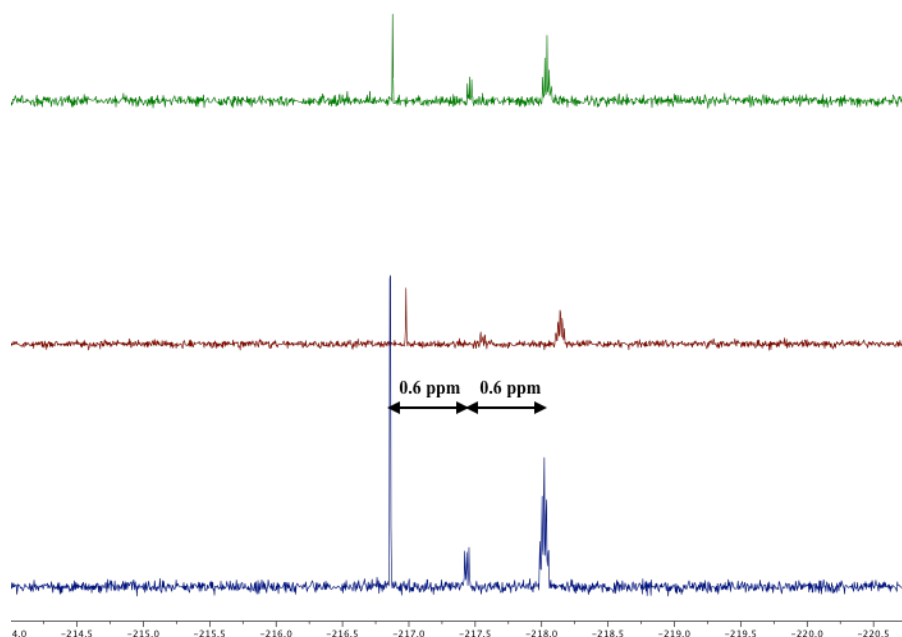
All three assays were incubated following standard conditions, at 30 °C and 180 rpm, and were sonicated, centrifuged and freeze-dried at day 12 for work up and analysed by  $^{19}\text{F}\{^1\text{H}\}$  NMR spectroscopy. The results are shown in **Fig. 17**.



**Fig. 17**  $^{19}\text{F}\{^1\text{H}\}$  NMR spectra after  $[5',5'\text{-}^2\text{H}_2]\text{-5'-FDA}$  (**17a**) incubations. A) Inoculation at day 0, no KF. B) Inoculation at days 4, 5 and 6, no KF. C) Inoculation at days 4, 5 and 6, 0.2 mM KF.

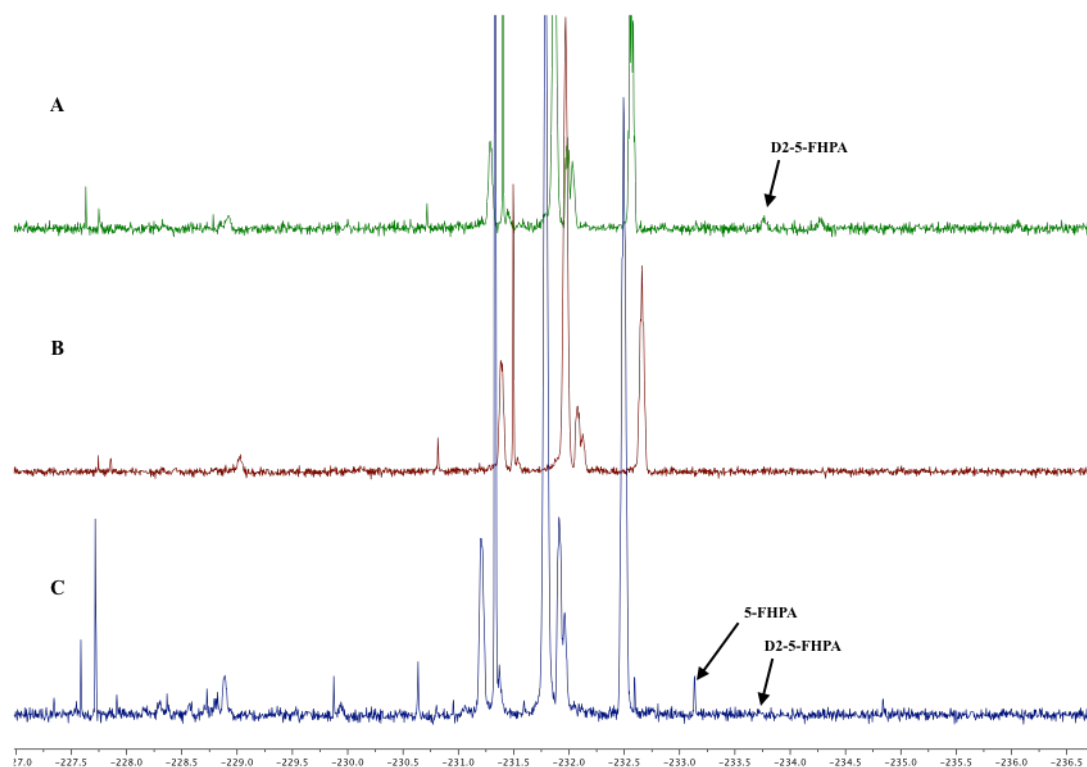
The formation of  $[^2\text{H}_2]\text{-fluoroacetate}$  **1a**, is obvious in the  $^{19}\text{F}$  NMR spectrum (**Fig. 17**), and shows that 5'-FDA is the source of fluorine for the first branch of the metabolic pathway. The three signals of fluoroacetate appreciated in the spectra appear at an approximate shift difference of 0.6 ppm from each other, corresponding to the populations of fluoroacetate that contain zero, one and two deuterium atoms, and they

are separated due to heavy atom upfield shifts, as previously established for deuterium labelled glycerols into FAc and 4-FT in *S. cattleya*.<sup>12</sup>



**Fig. 18** Expansion of  $^{19}\text{F}\{^1\text{H}\}$  NMR spectra for the fluoroacetate peaks observed in this study; showing heavy atom ( $^2\text{H}$ ) shifts (0.6 ppm) for populations of FAc (**1**),  $[\text{}^2\text{H}_1]\text{-FAc}$  (**1b**) and  $[\text{}^2\text{H}_2]\text{-FAc}$  (**1a**)

For confirmation of the source of fluoride in the second branch – to generate 5-FDR (**26**) and 5-FHPA (**2**) – it is necessary to enlarge the previous spectra (**Fig. 19**).



**Fig. 19**  $^{19}\text{F}\{^1\text{H}\}$  NMR spectra after  $[5',5'\text{-}^2\text{H}_2]\text{-5'-FDA}$  (**17a**) incubations (expansion). A) Inoculation at day 0, no KF. B) Inoculation at days 4, 5 and 6, no KF. C) Inoculation at days 4, 5 and 6, 0.2 mM KF.

The formation of deuterium incorporations into 5-FHPA (**2**) is clear in the third experiment supplemented with KF. The unlabelled 5-FHPA (**2**) has another peak 0.6 ppm upfield from this one, which suggests the formation of  $5\text{-}[^2\text{H}_1]\text{-5-FHPA}$ , and a further peak at approximately 1.2 ppm upfield from the unlabelled 5-FHPA is consistent with the formation of  $5\text{-}[^2\text{H}_2]\text{-5-FHPA}$  (**2a**).

Thus, it is concluded that 5'-FDA (**17**) is a precursor for 5-FHPA (**2**). The other unidentified compounds also seem to be similarly labelled, consistent with 5'-FDA being a common precursor. The inability to prepare  $5\text{-}[^2\text{H}_2]\text{-5-FDR}$  (**26a**) precluded this experiment; however, incubation with unlabelled 5-FDR **26** (Section 4.4.4, page 184) in CFEs did not generate FAc (**1**) nor 4-FT (**4**), indicating that it is uniquely formed only after the branch point.

## 4.8 Conclusions

In conclusion, the cell-free extracts of *S. sp.* MA37 proved to have poor fluorinase activity, when incubated with SAM and fluoride. Recombinant fluorinase was required to be added to generate a fluorometabolite profile.

The experiments that used 5'-FDA as a substrate led to the production, almost exclusively, of 5'-FDI; due to the dominance of an adenosine deaminase activity in the CFE. It proved important to have low 5'-FDA concentration, especially below 3 mM to generate a range of products.

An incubation experiment with 5'-FDI showed that this substrate was not metabolised in the CFE incubation, and it is concluded that it is a shunt metabolite.

CFE incubation with 5-FDR reproduced the original observation,<sup>1</sup> leading to the formation of numerous fluorometabolites. However, no FAc (**1**) or 4-FT (**4**) were generated, suggesting its role beyond the branch point.

Incubations of 5'-FDA (**17**) in live cultures led to the formation of a series of metabolites consistent with those observed in the metabolite profile of *S. sp.* MA37. Reproduction of this experiment with [5',5'-<sup>2</sup>H<sub>2</sub>]-5'-FDA (**17a**) allowed the confirmation that 5'-FDA is the universal source of all of the fluorometabolites for both branches of the metabolic pathway. Further experiments are needed for the identification of all of the isotopically labelled fluorinated peaks appearing in the spectra.

A synthesis of [5,5-<sup>2</sup>H<sub>2</sub>]-5-FDR (**26a**) was attempted, but yields were low, and its high water solubility precluded its isolation at a suitable level for a feeding experiment.

## 4.9 References

- 
- <sup>1</sup> L. Ma, A. Bartholomé, M. H. Tong, Z. Qin, Y. Yu, T. Shepherd, K. Kyeremeh, H. Deng, D. O'Hagan; *Chem. Sci.* **2015**, *6*, 1414.
- <sup>2</sup> F. Huang, S. F. Haydock, D. Spiteller, T. Mironenko, T.-L. Li, D. O'Hagan, P. F. Leadlay, J. B. Spencer; *Chem. Biol.* **2006**, *13*, 475.
- <sup>3</sup> A. S. Eustáquio, R. P. McGlinchey, Y. Liu, C. Hazzard, L. L. Beer, G. Florova, M. M. Alhamadsheh, A. Lechner, A. J. Kale, Y. Kobayashi, K. A. Reynolds, B. S. Moore; *PNAS*, **2009**, *106*, 12295.
- <sup>4</sup> H. Deng, D. O'Hagan, C. Schaffrath; *Nat. Prod. Rep.* **2004**, *21*, 773.
- <sup>5</sup> S. L. Cobb, *Doctoral Thesis*, **2006**.
- <sup>6</sup> D. O'Hagan, C. Schaffrath, S. L. Cobb, J. T. G. Hamilton, C. D. Murphy; *Nature*, **2002**, *416*, 279.
- <sup>7</sup> A. Bartholomé, J. E. Janso, U. Reilly, D. O'Hagan; *Org. Biomol. Chem.* **2017**, *15*, 61.
- <sup>8</sup> J. Epp, T. Widlanski; *J. Org. Chem.* **1999**, *64*, 293.
- <sup>9</sup> G. Tojo, M. Fernández; *Oxidation of Primary Alcohols to Carboxylic Acids*, Springer, **2007**.
- <sup>10</sup> R. N. Prasad, A. Fung, K. Tietje, H. H. Stein, H. D. Brondyk; *J. Med. Chem.* **1976**, *19*, 1180.
- <sup>11</sup> M. Szostak, D. Procter; *Org. Lett.*, **2014**, *16*, 5052.
- <sup>12</sup> J. Nieschalk, J. T. G. Hamilton, C. D. Murphy, D. B. Harper, D. O'Hagan; *Chem. Commun.* **1997**, *0*, 799.



## Chapter 5. Experimental

### 5.1 General Procedures and Methods

#### 5.1.1 Chemistry general procedures

All reactions were carried out in flame-dried glassware, which was cooled under argon gas. Air and/or humidity-sensitive reactions were carried out under an argon atmosphere. 'Room temperature' refers to conditions between 18 – 25 °C. Evaporation or concentration refers to the removal of solvent on a rotary evaporator attached to a diaphragm pump.

The reagents used were obtained from Alfa Aesar, Acros, Fisher, Fluorochem or Sigma Aldrich; and were used without any extra purification, unless stated otherwise. Anhydrous THF, DCM and Et<sub>2</sub>O were obtained from an MBRAUN SPS-800 Solvent Purification System by passage through two drying columns and were dispensed under an argon atmosphere. Anhydrous methanol and acetonitrile were distilled from calcium hydride in a recycling still, and were dispensed under an argon atmosphere.<sup>1</sup>

Analytical thin layer chromatography (TLC) was carried out on aluminium-backed silica gel 60 TLC plates with F<sub>254</sub> indicator (Merck KGaA). Visualisation of compounds was achieved under UV light (254 or 365 nm), or by staining with anisaldehyde-sulfuric acid solution or potassium permanganate solution, and subsequent thermal development. Preparative silica gel column chromatography was carried out with Geduran® Si 60 Silica Gel (40-63 µm, flash silica, Merck KGaA) under a positive pressure of compressed air, eluting with solvents as supplied.

HPLC analysis or semi-preparative purification were performed in a Shimadzu Prominence (SIL-20A HT autosampler, CL-20AT ternary pump, DGU-20A3R solvent degasser, SPD 20A UV detector and CBM-20A controller module), with the column specified in each case.

Proton (<sup>1</sup>H) and proton-decoupled nuclear magnetic resonance spectra (<sup>19</sup>F{<sup>1</sup>H}, <sup>13</sup>C{<sup>1</sup>H}) were recorded on Bruker Avance III 500 or Bruker Avance III HD 500 (500

MHz  $^1\text{H}$ , 476 MHz  $^{19}\text{F}$ , 126 MHz  $^{13}\text{C}$ ) spectrometers. Chemical shifts ( $\delta$ ) are expressed in ppm, and are quoted relative to the residual solvent signal. Proton coupling constants ( $J$ ) are given in Hertz (Hz), and quoted to the nearest 0.1 Hz. Identical coupling constants are averaged. When necessary, two-dimensional spectra (COSY, HSQC, HMBC) were used for the assignments of otherwise challenging signals.

Optical rotations were measured on a Perkin Elmer 341 polarimeter in a 1 dm cell at  $\lambda = 589$  nm (wavelength for the sodium D-line). Specific rotations are given in implied units of  $10^{-1} \text{ deg cm}^2 \text{ g}^{-1}$  and concentrations are given in g/100 mL.

High-resolution mass spectra were recorded using electrospray ionisation (ESI), on a Micromass LCT Spectrometer or on a ThermoFisher Excalibur Orbitrap Spectrometer from solutions of the analyte in methanol, acetonitrile or water. The Mass Spectrometry Service at the University of St Andrews was in charge of these results.

Melting points were measured on a Griffin MPA 350.BM2.5 apparatus or an Electrothermal IA9100 machine, and are uncorrected.

IR spectra were measured on an IRAffinity -1S Shimadzu Infra Red Spectrometer with diamond ATR attachment. Absorption maxima are reported in wavenumbers ( $\text{cm}^{-1}$ ).

X-ray diffraction data were collected by Prof. A. Slawin or Dr. David Cordes (University of St Andrews) on a Mercury 70 diffractometer which uses graphite monochromated Mo-K $\alpha$  radiation at 100 K. Structures were analysed on CrystalMaker. All the published structures have been deposited in CCDC. Please see **Appendix** for the complete account of the crystallographic data.

Frozen aqueous solutions were freeze-dried in a Christ Alpha 1-2 Plus freeze drier.

### 5.1.2 Biology general procedures

All glassware used for biological purposes was autoclaved in advance, and used inside a laminar flow cupboard (Gallenkamp). Plasticware was also sterilised by autoclaving, unless using new sterile material. Media was also sterilised by autoclaving.

Reagents used were purchased from Sigma-Aldrich, Alfa Aesar, Acros Organic or Fisher, with no further purification or treatment, unless otherwise stated. The synthetic and codon-optimised FdrC DNA plasmid dry was ordered from DNA2.0 (Gene 140635).<sup>2</sup> ADA enzyme from bovine spleen and PNP – recombinant, expressed in *E. coli* - are commercially available by Sigma-Aldrich. A second recombinant PNP was obtained from GSK. Phytase from wheat was commercially available by Sigma-Aldrich. NH enzyme was over-expressed and purified from a recombinant plasmid from *T. vivax* expressed in *E. coli*<sup>3</sup> and recombinant fluorinase was provided by Dr. Stephen Thompson (University of St. Andrews).<sup>4</sup> These enzymes and plasmids were stored at -20 to -80 °C.

Glycerol stocks of *Streptomyces* sp. MA37 were kindly provided by Dr. Long Ma (University of St. Andrews), and Dr. Hai Deng (University of Aberdeen) generously provided a Petri plate of *S. sp.* MA37.

All biological samples were flash-frozen and stored at -80 °C in a New Brunswick Scientific -80 °C freezer, and thawed in ice or at 4 °C before use.

Stock solutions of the substrates and cofactors used in the CFE reactions were made; the concentration of these solutions was of 10 mM, except for KF, which was 50 mM.

Incubations of less than 1.5 mL were carried out in a BTD Dry Block Heating System for microtubes (Grant), while bigger volumes or incubations requiring shaking were carried out into an I26 Incubator Shaker (New Brunswick Scientific).

Cell lysis was achieved in most cases by sonication with a Vibra Cell apparatus (Sonics & Materials Inc.), or performed in a one-shot cell disrupter (Constant Systems). Centrifugation was carried out on a Beckmann Avanti Centrifuge, at the conditions specified for each case.

Sodium dodecylsulfate polyacrylamide gel electrophoresis (SDS-PAGE) was run using NuPAGE<sup>TM</sup> 4-12% Bis Tris gels (Novex) in MOPS SDS running buffer, in a gel tank (XCell Sure Lock<sup>TM</sup>, Invitrogen, Novex Mini-Cell) connected to a Bio-Rad Power Pac 300 operating at a constant current of 125 mA for 50 minutes (Amersham Pharmacia biotech EPS 301). Protein samples for SDS-PAGE were prepared as follows, depending on their solubility. Soluble samples were prepared as a mixture of the corresponding solution (10  $\mu$ L), SDS buffer (7.5  $\mu$ L) and the lysis buffer used as specified for each case (20  $\mu$ L). Insoluble samples were prepared taking a pinhead amount of the solid sample, and mixing it with urea solution (50  $\mu$ L, 8 M) and SDS buffer (12.5  $\mu$ L). The gels were stained by soaking in Instant Blue<sup>TM</sup> reagent overnight.

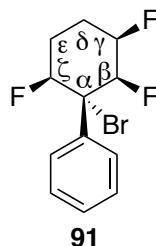
Optical density was measured at 600 nm in a WPA cell density meter. Protein concentrations, expressed in  $\text{mg mL}^{-1}$ , were determined using a Nanodrop ND-1000 spectrophotometer at 280 nm (Thermo Fisher Scientific). The extinction coefficients used were calculated by introducing the corresponding protein sequence into the ProtParam tool on the ExPASy Proteomics server ([www.expasy.org](http://www.expasy.org)).

Protein identification was analysed by an AB Sciex MALDI-TOF instrument, samples not giving conclusive data in this instrument were analysed by nanoLC-ESI MSMS using the Q-STAR XL. Whole protein mass spectroscopy was performed on the AB Sciex 4800 MALDI-TOF instrument.

Analytical analysis was carried out in the same way than for the chemical samples, as previously described.

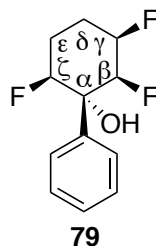
## 5.2 Synthetic preparation of compounds

### 5.2.1 Racemic ((1R,2R,3R,6S)-1-bromo-2,3,6-trifluorocyclohexyl)benzene (**91**)



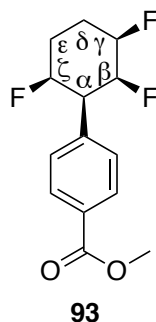
Racemic all-*cis* (2,3,6-trifluorocyclohexyl)benzene (**59**, 0.200 g, 0.93 mmol, 1 equiv.) was dissolved in CCl<sub>4</sub> (2 mL). NBS (210 mg, 1.12 mmol, 1.5 equiv.) was added to the solution. CH<sub>3</sub>CN (0.1 mL) was added for better solubilisation of **59**. The reaction mixture was heated at 90 °C for 48 h. The mixture was concentrated under pressure and redissolved in water (10 mL). The aqueous phase was extracted with Et<sub>2</sub>O (3 x 10 mL). The combined organic phases were dried over Na<sub>2</sub>SO<sub>4</sub> anhydrous, filtered and concentrated under reduced pressure. This procedure gave product **91** in quantitative yield and no further purification was needed. <sup>1</sup>H NMR (500 MHz, chloroform-*d*) δ<sub>H</sub> 7.68 (2H, dq, *J* = 8.3, 1.2 Hz, *H*-Ar), 7.44 (2H, m, *H*-Ar), 7.37 (1H, m, *H*-Ar), 5.74 (1H, dddt, *J* = 49.3, 8.3, 2.6, 1.2 Hz, *H*-γ), 5.5 (2H, m, *H*-β, *H*-ζ), 2.46 (3H, m, *H*-δ<sub>eq</sub>, *H*-ε), 2.01 (1H, dq, *J* = 8.3, 5.9, 4.2 Hz, *H*-δ<sub>ax</sub>) <sup>19</sup>F{<sup>1</sup>H} NMR (470 MHz, chloroform-*d*) δ<sub>F</sub> -189.8 (dd, *J* = 14.7, 18.5 Hz), -188.8 (d, *J* = 14.7 Hz), -171.8 (d, *J* = 18.5 Hz) <sup>13</sup>C NMR (126 MHz, chloroform-*d*) δ<sub>C</sub> 138.0 (*C*-Ar), 129.0 (d, *J* = 15.5 Hz), 126.3 (*C*-Ar), 90.4 – 86.5 (m, *C*-γ), 25.0 (dd, *J* = 22.4, 12.1 Hz, *C*-ε), 20.1 (dd, *J* = 20.4, 4.0 Hz, *C*-δ) HRMS (ESI<sup>+</sup>) *m/z* calc. for C<sub>12</sub>H<sub>12</sub>BrF<sub>3</sub>Na [M+Na]<sup>+</sup> 314.9972, found 314.9956.

### 5.2.2 Racemic (1*R*,2*S*,3*R*,6*S*)-2,3,6-trifluoro-1-phenylcyclohexan-1-ol (**79**)



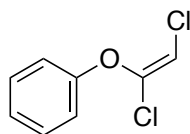
(1-Bromo-2,3,6-trifluorocyclohexyl)benzene (**91**, 40 mg, 0.14 mmol, 1 equiv.) was dissolved in DMF (2 mL). NaN<sub>3</sub> (25 mg, 0.30 mmol, 2 equiv.) was added and the mixture was heated to 80 °C for 18 h. The reaction was cooled to r. t. and quenched with water (50 mL). The aqueous phase was extracted with Et<sub>2</sub>O (3 x 50 mL). The combined organic phases were washed with brine (3 x 50 mL), dried over Na<sub>2</sub>SO<sub>4</sub> anhydrous, filtered and concentrated under reduced pressure. Further purification was carried out by column chromatography (100% petroleum ether), yielding hydroxylated **79** as a yellow oily product in 31% yield. <sup>1</sup>H NMR (500 MHz, chloroform-*d*) δ<sub>H</sub> 7.72 (2H, dd, *J* = 7.6, 1.7 Hz), 7.46 (2H, dd, *J* = 8.4, 6.6 Hz), 7.40 (1H, m), 5.17 (2H, m), 4.90 (dd, *J* = 45.8, 3.2 Hz, 1H), 2.22 (3H, m), 1.95 (1H, m) <sup>19</sup>F{<sup>1</sup>H} NMR (470 MHz, chloroform-*d*) δ<sub>F</sub> -189.1 (F, d, *J* = 23.0 Hz), -191.6 (F, m), -207.7 (F, m) <sup>13</sup>C NMR (126 MHz, chloroform-*d*) δ<sub>C</sub> 140.8, 127.3 (m), 90.0 (m), 73.8, 24.7 (dd, *J* = 21.8, 12.3 Hz), 20.0 (dd, *J* = 20.3, 4.4 Hz) HMRS (ESI<sup>+</sup>) *m/z* calc. for C<sub>12</sub>H<sub>13</sub>OF<sub>3</sub>Na [M+Na]<sup>+</sup> 253.0816, found 253.0812. NMR signals were assigned for **79** in the metabolic assays (Section 5.3.2.2).

### 5.2.3 Racemic methyl 4-((1*S*,2*S*,3*R*,6*S*)-2,3,6-trifluorocyclohexyl)benzoate (**93**)



4-(2,3,6-trifluorocyclohexyl)benzoic acid **60** (5 mg, 1 equiv.) was dissolved in a mixture CH<sub>3</sub>CN (200  $\mu$ L) and methanol (150  $\mu$ L). H<sub>2</sub>SO<sub>4</sub> (conc., 50  $\mu$ L) was added dropwise. The mixture was stirred at 80 °C for 16 h. The reaction mixture was cooled and quenched with Na<sub>2</sub>CO<sub>3</sub> (20% sol., added dropwise until pH 7.0). The neutral solution was extracted with DCM (3 x 1 mL). The combined organic phases were washed with water (3 x 3 mL), dried over anhydrous Na<sub>2</sub>SO<sub>4</sub>, filtered and under reduced pressure, resulting in methyl ester **93** in quantitative yield. No further purification was required. The enantiomerically enriched **60** recovered from the incubation with *C. elegans* was esterified following the same procedure, and the characterisation matches the racemic data. <sup>1</sup>H NMR (500 MHz, chloroform-*d*)  $\delta_{\text{H}}$  8.03 (d,  $J$  = 8.4 Hz, 2H, *H*-Ar), 7.60 – 7.52 (m, 2H, *H*-Ar), 5.10 (dd,  $J$  = 51.6, 8.6 Hz, 1H, *H*- $\gamma$ ), 4.88 (dt,  $J$  = 47.7, 2.9 Hz, 1H, *H*- $\beta$ ), 4.74 – 4.54 (m, 1H, *H*- $\zeta$ ), 3.92 (s, 3H, OCH<sub>3</sub>), 2.76 (t,  $J$  = 37.4 Hz, 1H, *H*- $\alpha$ ), 2.51 – 2.29 (m, 2H, *H*- $\epsilon$ ), 2.00 (dt,  $J$  = 13.2, 4.0 Hz, 1H, *H*- $\delta_{\text{eq}}$ ), 1.83 – 1.62 (m, 1H, *H*- $\delta_{\text{ax}}$ ). <sup>19</sup>F{<sup>1</sup>H} NMR (470 MHz, CDCl<sub>3</sub>)  $\delta_{\text{F}}$  -183.0, -191.0, -209.5 (dd,  $J$  = 26.8, 13.2 Hz). <sup>13</sup>C NMR (126 MHz, CDCl<sub>3</sub>)  $\delta_{\text{C}}$  166.9 (CO), 142.4 (*C*-Ar), 129.6 (d,  $J$  = 70.0 Hz, *C*- $\zeta$ ), 92.4 – 86.9 (m, *C*-Ar), 52.2, 48.4 (d,  $J$  = 5.4 Hz, *C*- $\alpha$ ), 29.7 (OCH<sub>3</sub>), 28.5 (dd,  $J$  = 22.5, 12.1 Hz, *C*- $\epsilon$ ), 20.4 (dd,  $J$  = 20.4, 3.6 Hz, *C*- $\delta$ ). HMRS (ESI<sup>+</sup>)  $m/z$  calc. for C<sub>14</sub>H<sub>16</sub>O<sub>2</sub>F<sub>3</sub> [M+H]<sup>+</sup> 273.1099, found 273.1098.

#### 5.2.4 (*E*)-((1,2-dichlorovinyl)oxy)benzene (**108**)

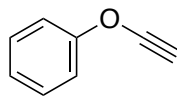


**108**

A solution of phenol **107** (5.20 g, 53 mmol) and ground sodium hydroxide (2.13 g, 53.0 mmol) in DMSO (50 mL) was stirred at room temperature for 2 h. 1,1,2-Trichloroethylene (4.80 mL, 53.0 mmol) is added, and the reaction mixture is stirred at r.t. for 18 h. The reaction mixture was quenched with addition of water (200 mL) and the aqueous phase was extracted thrice with CH<sub>2</sub>Cl<sub>2</sub> (3 x 150 mL). The combined organic phases were washed with brine, dried over MgSO<sub>4</sub>, filtered and evaporated under reduced pressure. Flash column chromatography in petroleum ether of the crude yields to the formation the title compound **108** (4.17 g, 22 mmol, 41% yield) as a colourless oil. <sup>1</sup>H NMR (500 MHz, chloroform-*d*): δ<sub>H</sub> 7.45 – 7.33 (m, 2H, *H*-Ar), 7.22 – 7.15 (m, 1H, *H*-Ar), 7.8 (dt, *J* = 7.8, 1.1 Hz, 2H, *H*-Ar), 5.97 (s, 1H, CClH). The data were in good agreement with the literature values.<sup>5</sup>



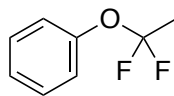
### 5.2.5 (Ethynyloxy)benzene (**109**)



**109**

A solution of (*E*)-1-(1,2-dichlorovinyl)oxybenzene **108** (2.17 g, 11.5 mmol) in Et<sub>2</sub>O (120 mL) was cooled to -78 °C and stirred for 30 min. *n*BuLi (18 mL, 46.00 mmol, 2.5 M) was added dropwise and stirred at -78 °C for 1 h. The mixture is left to reach -40 °C over the course of 1 h, and stirred at -40 °C for 2 h. The reaction mixture was quenched by addition of water (100 mL) at cold temperature. The organic phase was separated, and the aqueous phase extracted thrice with Et<sub>2</sub>O (3 x 100 mL). The combined organic phases were washed successively with a saturated solution of ammonium chloride and brine, dried over MgSO<sub>4</sub> and evaporated under reduced pressure. No further purification was required. Ethynyl(oxy)benzene **109** was obtained as a brown oil (1.17 g, 9.90 mmol, 86 % yield). <sup>1</sup>H NMR (500 MHz, chloroform-*d*): δ<sub>H</sub> 7.42 – 7.34 (m, 2H, *H*-Ar), 7.30 (dt, *J* = 7.9, 1.1 Hz, 2H, *H*-Ar), 7.22 – 7.12 (m, 1H, *H*-Ar), 2.09 (s, 1H, CCH). The data were in good agreement with the literature values.<sup>5</sup>

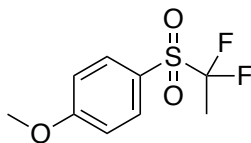
### 5.2.6 (1,1-Difluoroethoxy)benzene (**101**)



**101**

Ethynyl(oxy)benzene **109** (1 g, 8.47 mmol) was added into a 50 mL Teflon round-bottom flask and dissolved into  $\text{CH}_2\text{Cl}_2$  (5 mL). The solution was cooled to  $-78\text{ }^\circ\text{C}$  and stirred for 15 min.  $\text{HF}\cdot\text{py}$  (6 mL, 67.8 mmol) was added to the flask and the reaction mixture was stirred for 18 h and left to reach room temperature. The reaction was quenched by slow addition of the solution into a saturated solution of  $\text{NaHCO}_3$  (150 mL). The resulting mixture was extracted thrice with  $\text{CH}_2\text{Cl}_2$  (3 x 200 mL). The combined organic layers were washed with brine, dried over  $\text{MgSO}_4$  and evaporated under reduced pressure. The residue was further purified by washing with  $\text{CuSO}_4$  (7%, 25 mL) and extracting with  $\text{Et}_2\text{O}$  (3 x 25 mL), which afforded the pure product **101** as a colourless oil with a 75% conversion.  $^1\text{H NMR}$  (500 MHz, chloroform-*d*):  $\delta_{\text{H}}$  7.36 – 7.30 (m, 2H, *H*-Ar), 7.22 – 7.15 (m, 3H, *H*-Ar), 1.92 (t,  $J = 13.3\text{ Hz}$ , 3H,  $\text{CF}_2\text{CH}_3$ )  $^{19}\text{F NMR}$  (471 MHz, chloroform-*d*):  $\delta_{\text{F}}$  -64.4 (q,  $J = 13.2\text{ Hz}$ , 2F)  $^{13}\text{C NMR}$  (126 MHz, chloroform-*d*):  $\delta_{\text{C}}$  129.3 (*C*-Ar), 125.4 (*C*-Ar), 124.0 (t,  $J = 261.5\text{ Hz}$ ,  $\text{CF}_2\text{CH}_3$ ), 121.8 (*C*-Ar), 22.8 (t,  $J = 32.0\text{ Hz}$ ,  $\text{CF}_2\text{CH}_3$ ) **HMRS** (CI)  $\text{C}_8\text{H}_9\text{OF}_2$   $[\text{M}+\text{H}]^+$  calculated for 159.0621, found 159.0626.

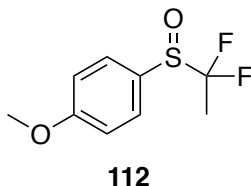
### 5.2.7 1-((1,1-Difluoroethyl)sulfonyl)-4-methoxybenzene (**115**)



**115**

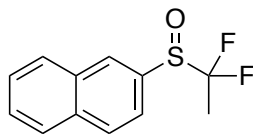
(1,1-Difluoroethyl)(4-methoxyphenyl)sulfane **104** (5 mg, 0.025 mmol) was added to a round bottom flask containing a stirring bar, and dissolved in CH<sub>2</sub>Cl<sub>2</sub> (1.5 mL). *m*CPBA was added to the solution (21 mg, 0.122 mmol), and the mixture was stirred at r.t. overnight. The reaction was quenched by addition of a saturated solution of NaHCO<sub>3</sub>. The aqueous phase was extracted with CH<sub>2</sub>Cl<sub>2</sub> (3 x 3 mL). The combined organic phases were combined, dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and concentrated under reduced pressure, yielding to **115** with 100% conversion. Further purification was carried out by column chromatography, starting with 100% petroleum ether, followed by 15% EtOAc in petroleum ether, affording **115** in quantitative yield. <sup>1</sup>H NMR (500 MHz, chloroform-*d*): δ<sub>H</sub> 7.84 (d, *J* = 8.6 Hz, 2H, *H*-Ar), 7.02 (d, *J* = 8.6 Hz, 2H, *H*-Ar), 3.91 (s, 3H, OCH<sub>3</sub>), 2.02 (t, *J* = 18.3 Hz, 3H, CF<sub>2</sub>CH<sub>3</sub>) <sup>19</sup>F NMR (471 MHz, chloroform-*d*): δ<sub>F</sub> -97.3 (s) <sup>13</sup>C NMR (126 MHz, chloroform-*d*): δ<sub>C</sub> 165.2, 133.1, 122.9, 114.7, 55.8 (OCH<sub>3</sub>), 16.6 (t, *J* = 22.2 Hz, CF<sub>2</sub>CH<sub>3</sub>) HMRS (ESI) C<sub>9</sub>H<sub>11</sub>F<sub>2</sub>O<sub>3</sub>S [M+H]<sup>+</sup> calculated for 236.0310, found 236.0390; [M+Na]<sup>+</sup> calculated for 259.0211, found 259.0216.

### 5.2.8 Racemic 1-((1,1-difluoroethyl)sulfinyl)-4-methoxybenzene (**112**)



(1,1-Difluoroethyl)(4-methoxyphenyl)sulfane **104** (5 mg, 0.025 mmol) was added to a round bottom flask with a stirring bar and dissolved in a mixture of DCM (3 mL) and methanol (0.3 mL). The solution was stirred at room temperature until homogenisation (5 min).  $\text{AlCl}_3$  (1.8 mg, 0.012 mmol) was added, and the solution stirred for 5 min, prior to the addition of BAIB (7.3 mg, 0.025 mmol). The reaction was left to stir overnight. After 16 h, the solvents were evaporated under reduced pressure. The remaining mixture showed the formation of **112** with 66% conversion from the starting material **104**. Further purification was achieved by reverse-phase HPLC in a Phenomenex Luna SP column, with 60:40 AcCN:water (supplemented with 0.05% TFA) at a flow rate of 1 mL/min. The product **112** was isolated at  $t_{\text{ret}} = 24$  min, which was consistent with the metabolic experiment's data.  $^1\text{H NMR}$  (500 MHz, chloroform-*d*):  $\delta_{\text{H}}$  7.65 (d,  $J = 8.9$  Hz, 2H, *H*-Ar), 7.08 (d,  $J = 8.9$  Hz, 2H, *H*-Ar), 3.89 (s, 1H,  $\text{OCH}_3$ ), 1.81 (t,  $J = 18.4$  Hz, 1H,  $\text{CF}_2\text{CH}_3$ )  $^{19}\text{F NMR}$  (471 MHz, chloroform-*d*):  $\delta_{\text{F}}$  -93.4 (d,  $J = 225.1$  Hz), -97.1 (d,  $J = 225.1$  Hz). Data is consistent with the full characterisation carried out for **112** from the metabolic assays (Section 5.3.3.1).

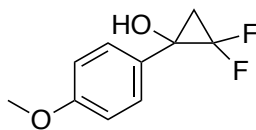
### 5.2.9 Racemic 2-((1,1-difluoroethyl)sulfinyl)naphthalene (**116**)



**116**

(1,1-Difluoroethyl)(naphthalene-2-yl) sulfane **105** (5 mg, 0.022 mmol) was added to a round bottom flask with a stirring bar and dissolved in a mixture of DCM (3 mL) and methanol (0.3 mL). The solution was stirred at room temperature until homogenisation (5 min).  $\text{AlCl}_3$  (1.5 mg, 0.011 mmol) was added, and the solution stirred for 5 min, prior to the addition of BAIB (11.1 mg, 0.022 mmol). The reaction was left to stir overnight. After 16 h, the solvents were evaporated under reduced pressure. Further purification was achieved by reverse-phase HPLC in a Phenomenex Luna SP column, with 60:40 AcCN:water (supplemented with 0.05% TFA) at a flow rate of 1 mL/min, which afforded **116** in 30% yield. The product **116** was isolated at  $t_{\text{ret}} = 37$  min, which was consistent with the metabolic experiments' data.  **$^1\text{H}$  NMR** (500 MHz, chloroform-*d*):  $\delta_{\text{H}}$  8.28 (s, 1H, *H*-Ar), 8.01 (d,  $J = 8.7$  Hz, 1H, *H*-Ar), 7.99 – 7.92 (m, 2H, *H*-Ar), 7.69 (ddt,  $J = 8.7, 2.6, 1.3$  Hz, 1H, *H*-Ar), 7.67 – 7.60 (m, 2H, *H*-Ar), 1.77 (t,  $J = 18.5$  Hz, 3H,  $\text{CF}_2\text{CH}_3$ )  **$^{19}\text{F}$  NMR** (471 MHz, chloroform-*d*):  $\delta_{\text{F}}$  -92.9 (d,  $J = 227.0$  Hz), -96.0 (d,  $J = 227.0$  Hz);  **$^{13}\text{C}$  NMR** (126 MHz, chloroform-*d*)  $\delta_{\text{C}}$  145.9 (*C*-Ar, visible in HMBC), 135.1 (s, *C*-Ar), 133.6 (t,  $J = 218.8$  Hz,  $\text{CF}_2$ ), 129.3 (s, *C*-Ar), 128.8 (s, *C*-Ar), 128.5 (s, *C*-Ar), 128.1 (s, *C*-Ar), 127.5 (s, *C*-Ar), 126.8 (s, *C*-Ar), 121.0 (s, *C*-Ar), 111.7 (*C*-Ar, visible in HMBC), 16.5 (t,  $J = 22.1$  Hz,  $\text{CF}_2\text{CH}_3$ ); **HRMS** ( $\text{ESI}^+$ )  $m/z$  calc. for  $\text{C}_{12}\text{H}_{11}\text{OF}_2\text{S}$   $[\text{M}+\text{H}]^+$  241.0420, found 241.0491.

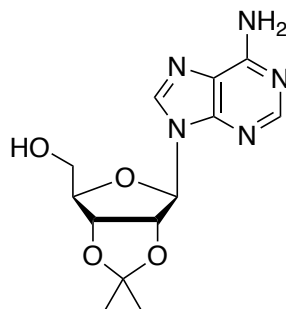
#### 5.2.10 2,2-Difluoro-1-(4-methoxyphenyl)cyclopropan-1-ol (**152**)



**152**

1-Methoxy-4-(1,2,2-trifluorocyclopropyl)benzene (**126**, 10 mg, 0.050 mmol) was added into a round bottom flask containing a stirring bar, and put under vacuum for 15 min. to ensure dryness. **126** was dissolved in DCM (2 mL), and the solution was cooled to 0 °C by placement in an iced water bath. BBr<sub>3</sub> was added (50 μL, 0.500 mmol), and the reaction was stirred until completion (2 h., 90% conversion). The reaction was quenched by addition of water (2 mL), and the pH adjusted to 7 by dropwise addition of a saturated NaHCO<sub>3</sub> solution. The aqueous phase was extracted with DCM (3 x 5 mL); and the combined organic phases were dried over anhydrous Na<sub>2</sub>SO<sub>4</sub>, filtered and concentrated under reduced pressure. Further purification of the reaction was carried out by column chromatography, using the following conditions: 1.) 100% petroleum ether. 2.) 15% EtOAc in petroleum ether. 3.) 30% EtOAc in petroleum ether. 4.) 50 % EtOAc in petroleum ether. Column chromatography afforded **152** as an oil. <sup>1</sup>H NMR (500 MHz, chloroform-*d*) δ<sub>H</sub> 7.40 (2H, d, *J* = 8.3 Hz, *H*-Ar), 6.89 (2H, d, *J* = 8.3 Hz, *H*-Ar), 3.82 (3H, s, OCH<sub>3</sub>), 2.20 (1H, ddd, *J* = 13.4, 9.6, 4.5 Hz, CF<sub>2</sub>CH<sub>2</sub>), 2.06 (1H, td, *J* = 9.6, 4.5 Hz, CF<sub>2</sub>CH<sub>2</sub>); <sup>19</sup>F NMR (471 MHz, CDCl<sub>3</sub>): δ<sub>F</sub> -127.0 (d, *J* = 147.0 Hz), -132.3 (d, *J* = 147.0 Hz); <sup>13</sup>C NMR (126 MHz, chloroform-*d*) δ<sub>C</sub> 160.1 (s, C-4), 130.7 (s, C-3), 128.3 (s, C-1), 114.3 (s, C-2), 109.2 (t, *J* = 293.6 Hz, CF<sub>2</sub>), 55.4 (s, OCH<sub>3</sub>), 27.3 (t, *J* = 10.6 Hz, CH<sub>2</sub>); HRMS (ESI) *m/z* calc. for C<sub>10</sub>H<sub>9</sub>O<sub>2</sub>F<sub>2</sub> [M-H]<sup>-</sup> 199.0576, found 199.0574.

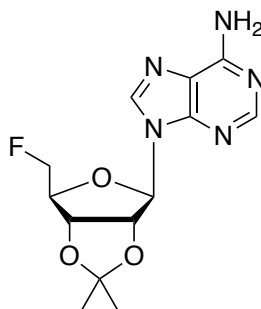
### 5.2.11 2',3'-*O*-Isopropylideneadenosine (**155**)



**155**

2,2-Dimethoxypropane (23 mL, 187 mmol, 10 eq.) and HClO<sub>4</sub> (2 mL, 1.5 eq., 70% aqueous) were added dropwise to a suspension of adenosine **154** (5 g, 18.7 mmol, 1 eq.) in acetone (160 mL) at 0 °C. The mixture was stirred at 0 °C for 6 h and left to warm to room temperature for 18 h. The reaction was quenched by addition of aqueous saturated NaHCO<sub>3</sub> solution (12 mL). The solvent was removed under reduced pressure, and the crude mixture was adsorbed onto silica gel. Purification of the residue over silica gel, eluting with petroleum ether and acetone (20:80), afforded **155** (3.6 g, 63%) as a white solid: **mp** 208-210 °C (MeOH); **[α]<sub>D</sub>** = -102.0 (*c* = 0.05 in MeOH); **<sup>1</sup>H NMR** (500 MHz, methanol-*d*<sub>4</sub>) δ<sub>H</sub> 8.35 (1H, s, *H*-2), 8.22 (1H, s, *H*-8), 6.18 (1H, d, *J* = 3.5 Hz, *H*-1'), 5.30 (1H, dd, *J* = 6.1 Hz, 3.5 Hz, *H*-2'), 5.07 (1H, dd, *J* = 6.1 Hz, 2.3 Hz, *H*-3'), 4.40 (1H, ddd, *J* = 3.7 Hz, 3.7 Hz, 2.3 Hz, *H*-4'), 3.81 (1H, dd, *J* = 12.1 Hz, 3.7 Hz, *H*-5'), 3.74 (1H, dd, *J* = 12.1 Hz, 3.7 Hz, *H*-5'), 1.64 (3H, s, C(CH<sub>3</sub>)<sub>2</sub>), 1.40 (s, 3H, C(CH<sub>3</sub>)<sub>2</sub>); **<sup>13</sup>C NMR** (126 MHz, methanol-*d*<sub>4</sub>) δ<sub>C</sub> 156.0 (*C*-5), 152.4 (*C*-2), 148.5 (*C*-4), 140.4 (*C*-8), 119.2 (*C*-6), 114.2 (C(CH<sub>3</sub>)<sub>2</sub>), 91.4 (*C*-1'), 86.5 (*C*-4'), 83.8 (*C*-2'), 81.5 (*C*-3'), 62.1 (*C*-5'), 26.2 (C(CH<sub>3</sub>)<sub>2</sub>), 24.2 (C(CH<sub>3</sub>)<sub>2</sub>); **HRMS** (ESI<sup>+</sup>) *m/z* calc. for C<sub>13</sub>H<sub>17</sub>N<sub>5</sub>O<sub>4</sub>H<sup>+</sup> [*M*+H]<sup>+</sup> 308.1353, found 308.1344.

#### 5.2.12 2',3'-*O*-Isopropylidene-5'-fluoro-5'-deoxyadenosine (**156**)

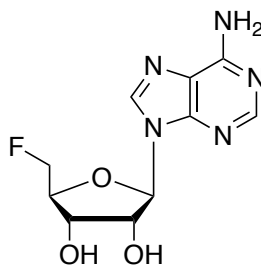


**156**

Tosyl fluoride (490 mg, 2.69 mmol) was set to dry under vacuo for 2h, prior to the addition of 2',3'-*O*-Isopropylideneadenosine (**155**) and a stirrer bar. The solid mixture was set to dry for further 30 min, before solution into THF (50 mL). TBAF (5.4 mL, 5.37 mmol) was added dropwise, and the mixture was set to reflux overnight at 60 °C. The solution was then allowed to cool to r.t., and was concentrated with silica gel under reduced pressure. Purification of **156** was carried out by dry-loading the resulting mixture into a chromatographic column, and eluting with 1:1 petroleum ether:acetone. **156** was recovered as a white solid in 75% yield. <sup>1</sup>H NMR (500 MHz, methanol-*d*<sub>4</sub>) δ<sub>H</sub> 8.25 (1H, s, *H*-2), 8.24 (1H, s, *H*-8), 6.27 (1H, d, *J* = 2.5 Hz, *H*-1'), 5.51-5.38 (1H, m, *H*-2'), 5.14 (1H, dd, *J* = 6.2 Hz, 3.1 Hz, *H*-3'), 4.73-4.55 (2H, m, *H*-5'), 4.54-4.45 (1H, m, *H*-4'), 1.63 (3H, s, C(CH<sub>3</sub>)<sub>2</sub>), 1.41 (s, 3H, C(CH<sub>3</sub>)<sub>2</sub>); <sup>19</sup>F NMR (471 MHz, methanol-*d*<sub>4</sub>): δ<sub>F</sub> – 230.4. Data matches literature reports.<sup>6</sup>



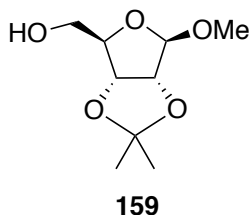
### 5.2.13 5'-Fluoro-5'-deoxyadenosine (5'-FDA, **17**)



**17**

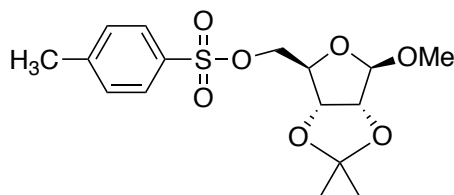
2',3'-*O*-Isopropylidene-5'-fluoro-5'-deoxyadenosine **156** (417 mg, 1.35 mmol) was added to a round bottom flask with a stirring bar, and dissolved in a mixture of THF (5 mL) and water (5 mL). Hydrated *p*-toluenesulfonic acid (369 mg, 1.62 mmol) was added, and the mixture was set to reflux at 63 °C. After 6 h, the reaction has gone to completion. THF was evaporated under reduced pressure, and saturated NaHCO<sub>3</sub> (50 mL) was added. The aqueous phase was extracted with EtOAc (3 x 100 mL), and the combined organic phases were dried over anhydrous Na<sub>2</sub>SO<sub>4</sub>, filtered and concentrated under vacuum. Further purification was carried out by dry-loading the resulting mixture into a chromatographic column, and eluting with 10% MeOH in DCM. Purification afforded **17** as a white solid in 39% yield. <sup>1</sup>H NMR (500 MHz, methanol-*d*<sub>4</sub>) δ<sub>H</sub> 8.24 (1H, s, *H*-2), 8.23 (1H, s, *H*-8), 6.09 (1H, d, *J* = 4.4 Hz, *H*-1'), 4.81-4.72 (1H, m, *H*-3'), 4.72 (2H, ddd, *J* = 47.0, 20.6, 3.1 Hz, *H*-5'), 4.64-4.59 (1H, m, *H*-4'), 4.26 (1H, ddt, *J* = 28.1, 5.7, 3.1 Hz, *H*-2'); <sup>19</sup>F NMR (471 MHz, deuterium oxide): δ<sub>F</sub> -231.3. Data matches literature reports.<sup>6,7</sup>

#### 5.2.14 2,3-*O*-Isopropylidene-1-methoxy ribose (**159**)



Ribose (**157**, 2.5 g, 16.6 mmol) was added into a round bottom flask containing a stirring bar and dissolved in acetone (200 mL). The suspension was cooled to 0 °C in an iced water bath. 2,2-Dimethoxypropane (5.1 mL, 41.5 mmol) was added, followed by dropwise addition of H<sub>2</sub>SO<sub>4</sub> (70%, 1.6 mL, 25 mmol). The reaction was stirred and warmed to r.t. overnight. The mixture is bright orange. Without quenching, MeOH (3.5 mL) was added. The reaction was stirred at r.t. for further 3 h. Then, the reaction was quenched at 0 °C by addition of a saturated solution of NaHCO<sub>3</sub> (10 mL), producing a suspension. The solid was filtered off, and the solution concentrated under reduced pressure. The residue was dissolved in Et<sub>2</sub>O (100 mL) and water (100 mL). The aqueous phase was extracted with Et<sub>2</sub>O (3 x 100 mL), and the combined organic phases were dried over anhydrous Na<sub>2</sub>SO<sub>4</sub>, filtered and concentrated under reduced pressure, giving **159** in 83% yield and a white powder. Further purification was not required. <sup>1</sup>H NMR (500 MHz, methanol-*d*<sub>4</sub>) δ<sub>H</sub> 4.92 (1H, s, *H*-1), 4.72 (1H, d, *J* = 6.0 Hz, *H*-2), 4.59 (1H, d, *J* = 6.0 Hz, *H*-3), 4.19 – 4.14 (1H, m, *H*-4), 3.60 – 3.50 (2H, m, *H*-5), 3.34 (3H, s, OCH<sub>3</sub>), 1.45 (3H, s, C(CH<sub>3</sub>)<sub>2</sub>), 1.33 (3H, s, C(CH<sub>3</sub>)<sub>2</sub>); <sup>13</sup>C NMR (126 MHz, methanol-*d*<sub>4</sub>) δ<sub>C</sub> 111.93 (C(CH<sub>3</sub>)<sub>2</sub>), 109.3 (C-1), 87.3 (C-4), 85.1 (C-3), 81.7 (C-2), 62.8 (C-5), 53.8 (OCH<sub>3</sub>), 25.3 (C(CH<sub>3</sub>)<sub>2</sub>), 23.6 (C(CH<sub>3</sub>)<sub>2</sub>). Experimental data matched previous literature data.<sup>8</sup>

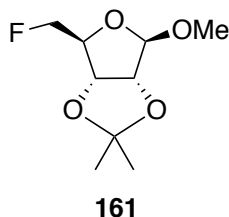
### 5.2.15 2,3-*O*-Isopropylidene-1-methoxy-5-tosyl ribose (**160**)



**160**

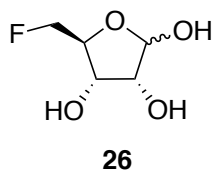
2,3-*O*-Isopropylidene-1-methoxy ribose (**159**, 2.4 g, 11.8 mmol) was added into a round bottom flask and dissolved in THF (200 mL). Tosyl fluoride (3 g, 17.6 mmol) was added to the solution, followed by dropwise addition of TBAF (35 mL, 35.2 mmol). The mixture was stirred under reflux (60 °C) overnight. Purification was carried out by dryloading the reaction's residue into a chromatographic column, and eluting with 10% EtOAc in petroleum ether, which afforded a crystalline white solid (**160**) in 63% yield. <sup>1</sup>H NMR (500 MHz, chloroform-*d*)  $\delta_{\text{H}}$  7.80 (2H, d,  $J = 8.3$  Hz, *H*-Ar), 7.37 (2H, d,  $J = 8.3$  Hz, *H*-Ar), 4.92 (1H, s, *H*-1), 4.59 (1H, dd,  $J = 6.0, 1.2$  Hz, *H*-2), 4.53 (1H, d,  $J = 6.0$  Hz, *H*-3), 4.31 (1H, td,  $J = 7.2, 1.1$  Hz, *H*-4), 4.01 (2H, dd,  $J = 7.2, 1.9$  Hz, *H*-5), 3.22 (3H, s, OCH<sub>3</sub>), 2.45 (3H, s, Ts-CH<sub>3</sub>), 1.44 (3H, s, C(CH<sub>3</sub>)<sub>2</sub>), 1.28 (3H, s, C(CH<sub>3</sub>)<sub>2</sub>). Experimental data matches previous literature data.<sup>8</sup>

### 5.2.16 5-Fluoro-5-deoxy-2,3-*O*-isopropylidene-1-methoxy ribose (**161**)



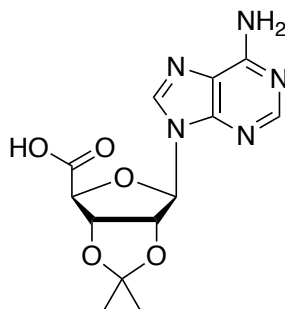
2,3-*O*-Isopropylidene-1-methoxy-5-tosyl ribose (**160**, 1.5 g, 4.2 mmol) was added to a round bottom flask containing a stirring bar and dissolved in AcCN (18 mL). TBAF (5 mL, 5 mmol) was added dropwise. The solution was stirred under reflux (80 °C) for 4 h. The solution was evaporated with silica gel under reduced pressure and dryloaded into a chromatographic column (eluted with 30:1 Petroleum ether:EtOAc), which gave **161** as a white solid in 80% yield. <sup>1</sup>H NMR (500 MHz, chloroform-*d*) δ<sub>H</sub> 4.99 (1H, d, *J* = 2.4 Hz, *H*-1), 4.71 – 4.68 (1H, m, *H*-2), 4.59 (1H, d, *J* = 6.0 Hz, *H*-3), 4.46 – 4.30 (3H, m, *H*-4, *H*-5), 3.33 (3H, s, OCH<sub>3</sub>), 1.49 (3H, s, C(CH<sub>3</sub>)<sub>2</sub>), 1.32 (3H, s, C(CH<sub>3</sub>)<sub>2</sub>). <sup>19</sup>F NMR (471 MHz, CDCl<sub>3</sub>) δ<sub>F</sub> -225.0 (s). Experimental data matches previous literature data.<sup>8</sup>

### 5.2.17 5-Fluoro-5-deoxy-2,3-*O*-isopropylidene-1-methoxy ribose (**5-FDR**, **26**)



5-Fluoro-5-deoxy-2,3-*O*-isopropylidene-1-methoxy ribose (**161**, 90 mg, 0.44 mmol) was dissolved in H<sub>2</sub>SO<sub>4</sub> (0.02 M, 1 mL, 0.02 mmol), and set under reflux (98 °C) for 1.5 h. The cooled solution was neutralised to pH = 7 by dropwise addition of a saturated solution of BaCO<sub>3</sub>. The suspension was filtered through a pad of celite and cotton, and the solution was freeze-dried over night. This afforded **26** as a colourless oil in 30% yield. <sup>1</sup>H NMR (500 MHz, deuterium oxide) δ<sub>H</sub> 5.32 (1H, d, *J* = 4.0 Hz, *H*-α anomer), 5.18 (1H, d, *J* = 1.7 Hz, *H*-β anomer), other signals could not be interpreted due to overlap. <sup>19</sup>F NMR (471 MHz, deuterium oxide) δ<sub>F</sub> -228.7 (s, β anomer), -231.0 (s, α anomer); <sup>13</sup>C NMR (126 MHz, methanol-*d*<sub>4</sub>) δ<sub>C</sub> 101.1 (*C*-1, β anomer), 96.5 (*C*-1, α anomer), 83.3 (d, *J* = 168.8 Hz, *C*-5, β anomer), 82.9 (d, *J* = 168.8 Hz, *C*-5, α anomer), 81.0 (d, *J* = 19.6 Hz, *C*-4, α anomer) 80.5 (d, *J* = 19.6 Hz, *C*-4, β anomer), 75.0 (*C*-2, β anomer), 70.6 (*C*-2, α anomer), 69.4 (d, *J* = 6.5 Hz, *C*-3, β anomer), 69.2 (d, *J* = 6.5 Hz, *C*-3, α anomer). Experimental data matches previous literature data.<sup>8</sup>

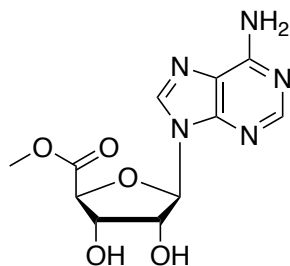
### 5.2.18 2',3'-*O*-Isopropylidene adenosine-5'-carboxylic acid (**167**)



**167**

BAIB (11.2 g, 34.2 mmol, 2.2 eq.) and TEMPO (0.5 g, 3.3 mmol, 0.2 eq.) were added to a solution of **155** (4.8 g, 15.6 mmol, 1 eq.) in a mixture of H<sub>2</sub>O:CH<sub>3</sub>CN (1:1, 40 mL). The mixture was stirred at room temperature for 18 h. The suspension was filtered, and the residue was washed with water (20 mL), acetone (20 mL) and ether (20 mL) successively. The residue was dried under vacuum, furnishing **167** (4.3 g, 85%) as a white solid: **mp** 260–265 °C (dec.); **<sup>1</sup>H NMR** (500 MHz, DMSO-*d*<sub>6</sub>) δ<sub>H</sub> 8.26 (1H, s, *H*-8), 8.09 (1H, s, *H*-2), 7.29 (2H, s, NH<sub>2</sub>), 6.33 (1H, s, *H*-1'), 5.54 (1H, dd, *J* = 6.0, 1.9 Hz, *H*-3'), 5.47 (1H, d, *J* = 6.0 Hz, *H*'-2), 4.69 (1H, d, *J* = 1.9 Hz, *H*-4'), 1.53 (3H, s, C(CH<sub>3</sub>)<sub>2</sub>), 1.36 (3H, s, C(CH<sub>3</sub>)<sub>2</sub>); **<sup>13</sup>C NMR** (126 MHz, DMSO-*d*<sub>6</sub>) δ<sub>C</sub> 171.2 (*C*-5'), 156.5 (*C*-2), 152.8 (*C*-5), 149.6 (*C*-4), 140.8 (*C*-8), 119.2 (*C*-6), 113.1 (C(CH<sub>3</sub>)<sub>2</sub>), 89.9 (*C*-1'), 85.9 (*C*-4'), 84.3 (*C*-3'), 83.9 (*C*-2'), 27.0 (C(CH<sub>3</sub>)<sub>2</sub>), 25.4 (C(CH<sub>3</sub>)<sub>2</sub>); ***m/z*** (ESI<sup>+</sup>) 322 ([*M*+*H*]<sup>+</sup>, 100%); **HRMS** (ESI<sup>+</sup>) *m/z* calc. for C<sub>13</sub>H<sub>15</sub>N<sub>5</sub>O<sub>4</sub>H<sup>+</sup> [*M*+*H*]<sup>+</sup> 322.1151, found 322.1136.

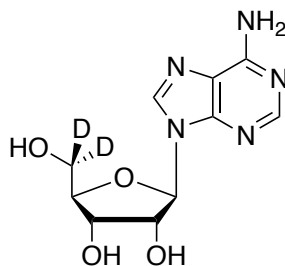
### 5.2.19 Adenosine-5'-carboxylic acid methyl ester (**169**)



**169**

Thionyl chloride (0.6 mL, 17.1 mmol, 5 eq.) was added dropwise to a suspension of **167** (1.1 g, 3.4 mmol, 1 eq.) in methanol (30 mL) that was stirred at 0°C for 30 min. The mixture stirred for 18 h, gradually warming to room temperature. The white suspension was concentrated and the residue was purified over silica gel, eluting with methanol and DCM (10:90), to afford **169** (0.9 g, 94%) as a white solid: **mp** 110-115 °C;  $[\alpha]_D^{25}$  = -110.0 ( $c$  = 0.07 in MeOH); **<sup>1</sup>H NMR** (500 MHz, methanol-*d*<sub>4</sub>)  $\delta_H$  8.86 (1H, s, H-8), 8.43 (1H, s, H-2), 6.30 (1H, d,  $J$  = 6.1 Hz, H-1'), 4.68 (1H, d,  $J$  = 2.7 Hz, H-4'), 4.64 (1H, dd,  $J$  = 6.1, 4.5 Hz, H-2'), 4.47 (1H, dd,  $J$  = 4.5, 2.7 Hz, H-3'), 3.37 (3H, s, OCH<sub>3</sub>); **<sup>13</sup>C NMR** (126 MHz, methanol-*d*<sub>4</sub>)  $\delta_C$  171.4 (C-5'), 150.6 (C-5), 149.0 (C-4), 144.3 (C-2), 142.4 (C-8), 118.8 (C-6), 88.5 (C-1'), 83.2 (C-4'), 75.3 (C-2'), 73.6 (C-3'), 51.8 (-OCH<sub>3</sub>); ***m/z*** (ESI<sup>+</sup>) 296 ([M+H]<sup>+</sup>, 100%); **HRMS** *m/z* (ESI<sup>+</sup>) calc. for C<sub>11</sub>H<sub>13</sub>N<sub>5</sub>O<sub>5</sub>H<sup>+</sup> [M+H]<sup>+</sup> 296.0989, found 296.0981.

### 5.2.20 5',5'-<sup>2</sup>H<sub>2</sub>-Adenosine (**154a**)

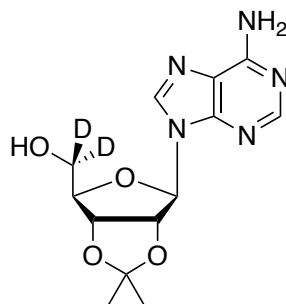


**154a**

NaBD<sub>4</sub> (160 mg, 4.1 mmol, 5 eq.) was added in small portions to a solution of **169** (240 mg, 0.8 mmol, 1 eq.) in ethanol (25 mL) at 0 °C. The mixture was stirred at 0 °C for 5 min and heated under reflux for 48 h. The mixture was concentrated to a powder, the residue was dissolved in water and the pH was adjusted to 6.5 by dropwise addition of HCl (1M). Concentration of the mixture afforded **154a** (152 mg, 69%) as a white powder: **mp** 170-175 °C (MeOH); **<sup>1</sup>H NMR** (500 MHz, methanol-*d*<sub>4</sub>) δ<sub>H</sub> 8.33 (1H, s, *H*-8), 8.20 (1H, s, *H*-2), 5.98 (1H, d, *J* = 6.4 Hz, *H*-1'), 4.76 (1H, dd, *J* = 6.4, 5.1 Hz, *H*-2'), 4.34 (1H, dd, *J* = 5.1, 2.6 Hz, *H*-3'), 4.18 (1H, d, *J* = 2.6 Hz, *H*-4'); **<sup>13</sup>C NMR** (126 MHz, methanol-*d*<sub>4</sub>) δ<sub>C</sub> 156.2 (*C*-5), 152.1 (*C*-2), 148.6 (*C*-4), 140.6 (*C*-8), 119.6 (*C*-6), 89.9 (*C*-1'), 86.7 (*C*-4'), 74.0 (*C*-2'), 71.3 (*C*-3'), 61.7 (*C*-5'); ***m/z*** (ESI<sup>+</sup>) 292 ([*M*+Na]<sup>+</sup>, 100%); **HRMS** (ESI<sup>+</sup>) *m/z* calc. for C<sub>10</sub>H<sub>11</sub>D<sub>2</sub>N<sub>5</sub>O<sub>4</sub>H<sup>+</sup> [*M*+H]<sup>+</sup> 270.1166, found 270.1164.



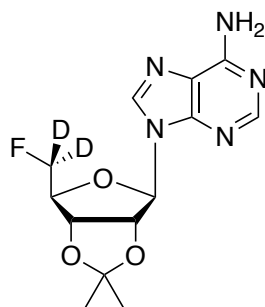
### 5.2.21 [5',5'-<sup>2</sup>H<sub>2</sub>]-2',3'-O-Isopropylideneadenosine (**155a**)



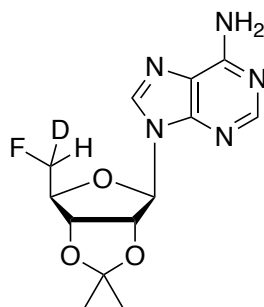
**155a**

2,2-dimethoxypropane (3.3 mL, 27.4 mmol, 10 eq.) and HClO<sub>4</sub> (330 μL, 1 mmol, 1.5 eq.) were added successively to a solution of **154a** (740 mg, 2.7 mmol, 1 eq.) in acetone (22.4 mL) at 0 °C. The mixture was stirred at 0 °C for 6 h, and stirred for 18 h, gradually warming to room temperature. The reaction was quenched with aqueous saturated NaHCO<sub>3</sub> solution (1.5 mL) and concentrated. The resultant white solid was purified over silica gel, eluting with petroleum ether and acetone (20:80), to yield **155a** (400 mg, 48% yield) as a white powder: **mp** 208-210 °C (MeOH); [ $\alpha$ ]<sub>D</sub> -96.4 (*c* = 0.06 in MeOH); <sup>1</sup>H NMR (500 MHz, methanol-*d*<sub>4</sub>) δ<sub>H</sub> 8.35 (1H, s, *H*-8), 8.22 (1H, s, *H*-2), 6.18 (1H, d, *J* = 3.5 Hz, *H*-1'), 5.30 (1H, dd, *J* = 6.1, 3.5 Hz, *H*-2'), 5.06 (1H, dd, *J* = 6.1, 2.3 Hz, *H*-3'), 4.39 (1H, d, *J* = 2.3 Hz, *H*-4'), 1.64 (3H, s, C(CH<sub>3</sub>)<sub>2</sub>), 1.40 (3H, s, C(CH<sub>3</sub>)<sub>2</sub>); <sup>13</sup>C NMR (126 MHz, methanol-*d*<sub>4</sub>) δ<sub>C</sub> 156.1 (*C*-5), 152.4 (*C*-2), 148.6 (*C*-4), 140.4 (*C*-8), 119.3 (*C*-6), 113.9 (C(CH<sub>3</sub>)<sub>2</sub>), 91.5 (*C*-1'), 86.6 (*C*-4'), 83.9 (*C*-2'), 81.6 (*C*-3'), 61.3 (*C*-5') 26.2 (C(CH<sub>3</sub>)<sub>2</sub>), 24.1 (C(CH<sub>3</sub>)<sub>2</sub>); *m/z* (ESI<sup>+</sup>) 332 ([M+Na]<sup>+</sup>, 100%); **HRMS** (ESI<sup>+</sup>) *m/z* calc. for C<sub>13</sub>H<sub>15</sub>D<sub>2</sub>N<sub>5</sub>O<sub>4</sub>H<sup>+</sup> [M+H]<sup>+</sup> 310.1479, found 310.1466.

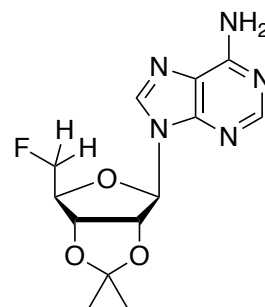
### 5.2.22 5'-Deoxy-5'-fluoro-5',5'-<sup>2</sup>H<sub>2</sub>-2',3'-*O*-isopropylideneadenosine (**156a**)



**156a**



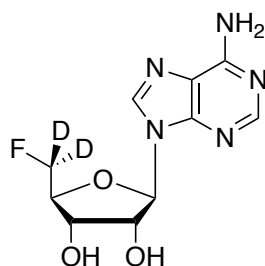
**156b**



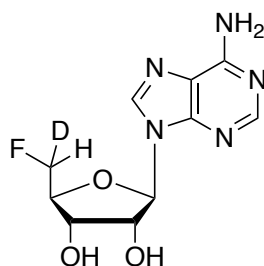
**156**

TsF (337 mg, 1.9 mmol, 1.5 eq.) and TBAF (4 mL, 3.9 mmol, 3 eq., 1.0 M solution in THF) were added to a solution of **155a** (400 mg, 1.3 mmol, 1 eq.) in THF (40 mL). The mixture was heated under reflux for 18 h. The solution was cooled, and concentrated and purified over silica gel, eluting with petroleum ether and acetone (50:50), to afford **156a**, **156b** and **156** (250 mg, 63%) as a colourless solid: **mp** 130–135 °C (MeOH);  $[\alpha]_D^{25}$  = -58.0 ( $c$  = 0.05 in MeOH);  $^1\text{H NMR}$  (500 MHz, methanol- $d_4$ )  $\delta_{\text{H}}$  8.24 (1H, s,  $H$ -8), 8.23 (1H, s,  $H$ -2), 6.26 (1H, d,  $J$  = 2.5 Hz,  $H$ -1'), 5.42 (1H, dd,  $J$  = 6.3 Hz, 2.5 Hz,  $H$ -2'), 5.13 (1H, dd,  $J$  6.3 Hz, 3.1 Hz,  $H$ -3'), 4.48 (1H, dd,  $J$  = 24.5, 3.1 Hz,  $H$ -4'), 1.65 (3H, s,  $\text{C}(\text{CH}_3)_2$ ), 1.47 (3H, s,  $\text{C}(\text{CH}_3)_2$ );  $^{19}\text{F}\{^1\text{H}\}$  NMR (470 MHz, methanol- $d_4$ )  $\delta_{\text{F}}$  -230.4 (**156**), -231.1 (**156b**), -231.7 (**156a**);  $^{13}\text{C NMR}$  (126 MHz, methanol- $d_4$ )  $\delta_{\text{C}}$  156.0 ( $\text{C}$ -5), 152.6 ( $\text{C}$ -2), 148.9 ( $\text{C}$ -4), 139.8 ( $\text{C}$ -8), 119.0 ( $\text{C}$ -6), 114.1 ( $\text{C}(\text{CH}_3)_2$ ), 90.4 ( $\text{C}$ -1'), 85.4 ( $\text{C}$ -2'), 84.2 ( $\text{C}$ -4'), 83.0 (m,  $\text{C}$ -5'), 80.7 ( $\text{C}$ -3'), 26.0 ( $\text{C}(\text{CH}_3)_2$ ), 24.1 ( $\text{C}(\text{CH}_3)_2$ );  $m/z$  ( $\text{ESI}^+$ ) 312 ( $[\text{M}+\text{H}]^+$ , 100%); **HRMS** ( $\text{ESI}^+$ )  $m/z$  calc. for  $\text{C}_{13}\text{H}_{14}\text{D}_2\text{FN}_5\text{O}_3\text{H}^+$   $[\text{M}+\text{H}]^+$  311.1444, found 311.1422.

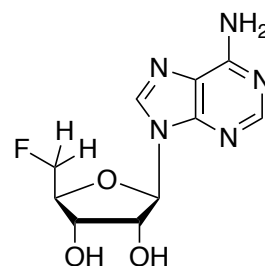
### 5.2.23 5'-Deoxy-5'-fluoro-5',5'-<sup>2</sup>H<sub>2</sub>-adenosine (17a)



**17a**



**17b**



**17**

*p*-TsOH·H<sub>2</sub>O (44 mg, 0.3 mmol, 1.2 eq.) was added to a solution of 5',5'-<sup>2</sup>H<sub>2</sub>-5'-deoxy-5'-fluoro-adenosine **156a** (80 mg, 0.3 mmol, 1 eq.) in a mixture of MeOH:H<sub>2</sub>O (1:1, 3 mL). The mixture was heated to 60 °C for 2 h and was cooled and concentrated. The residue was purified over silica gel, eluting with MeOH and DCM (10:90), to give **17a** (29 mg, 35%) as a white solid: **mp** 130-135 °C (H<sub>2</sub>O); [ $\alpha$ ]<sub>D</sub> = -25.1 (*c* = 0.18 in MeOH); <sup>1</sup>H NMR (500 MHz, methanol-*d*<sub>4</sub>)  $\delta$ <sub>H</sub> 8.38 (1H, s, *H*-8), 8.35 (1H, s, *H*-2), 6.13 (1H, d, *J* = 4.7 Hz, *H*-1'), 4.61 (1H, dd, *J* = 4.7, 1.2 Hz, *H*-2'), 4.42 (1H, dd, *J* = 5.1, 1.2 Hz, *H*-3'), 4.27 (1H, dd, *J* = 28.7, 5.1 Hz, *H*-4'); <sup>19</sup>F{<sup>1</sup>H} NMR (470 MHz, methanol-*d*<sub>4</sub>)  $\delta$ <sub>F</sub> -232.5 (**17**), -233.1 (**17b**), -233.7 (**17a**); <sup>13</sup>C NMR (126 MHz, methanol-*d*<sub>4</sub>)  $\delta$ <sub>C</sub> 155.9 (*C*-5), 152.7 (*C*-2), 149.1 (*C*-4), 139.3 (*C*-8), 119.0 (*C*-6), 88.5 (*C*-1'), 83.1 (*C*-4'), 81.1 (m, *C*-5'), 74.3 (*C*-2'), 69.5 (*C*-3'); *m/z* (ESI<sup>+</sup>) 272 ([M+H]<sup>+</sup>, 100%); **HRMS** (ESI<sup>+</sup>) *m/z* calc. for C<sub>10</sub>H<sub>10</sub>D<sub>2</sub>FN<sub>5</sub>O<sub>3</sub>H<sup>+</sup> [M+H]<sup>+</sup> 272.1123, found 272.1116.

### 5.3 Fluorometabolite production in *Cunninghamella elegans*

#### 5.3.1 General procedures for the works with *C. elegans*

##### 5.3.1.1 Preparation of liquid cultures of *C. elegans* and inoculation of the corresponding xenobiotics

Sterile Saboraud Dextrose (Sigma Aldrich) medium (45 mL) was inoculated with fungal homogenate (5 mL) at room temperature. Alternatively, inoculation was carried out directly from fungal gel plates into the SD liquid medium (50 mL). The cultures were left to grow for 72 h, at 28 °C with rotary agitation (150 rpm).

All xenobiotics were added dissolved in DMF (50 µL) to grown cultures (50 mL, 72 h or prior incubation), and left to incubate for extra 72 h at 28 °C and 150 rpm.

##### 5.3.1.2 Extraction and purification of the metabolites

After the incubation period, the xenobiotic-supplemented liquid cultures were centrifuged using a Sorvall apparatus at 4 °C. The supernatant was separated from the fungal debris and extracted with organic solvents; usually EtOAc (3×50 mL), unless otherwise stated. The organic phase was concentrated under reduced pressure, and the solid residues redissolved in EtOAc (1mL). The fungal cells were stored at –20 °C.

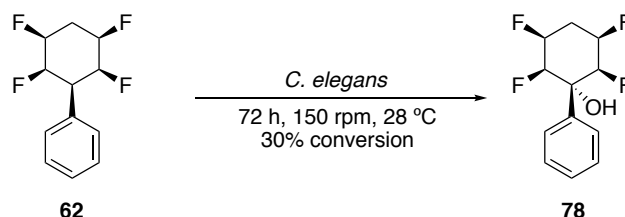
For the fluorocyclohexyl derivatives, an aliquot of the concentrate (100 µL) was added to MSTFA (50 µL) for TMS derivatisation of the hydroxyl groups (1 h, 100 °C) prior to analysis by GCMS. Further analysis of the extracts was done by <sup>1</sup>H and <sup>19</sup>F NMR.

Purification of the samples was mainly performed by reverse phase HPLC, using either a semi-preparative Kingsorb Phenomenex or a semi-preparative Luna Phenomenex. Usual conditions included 70:30 of AcCN:water, 60:40 AcCN:water, or 50:50 AcCN:water, as indicated for each compound. Both water and acetonitrile were supplemented with 0.05% TFA, unless otherwise expressed. Some of the metabolites required the use of column chromatography, as described in the corresponding experimental procedures. Analysis of the resulting metabolites and remaining starting materials was carried out by full NMR characterisation (<sup>1</sup>H, <sup>19</sup>F, <sup>13</sup>C, COSY, HSQC

and HMBC), and high-resolution accurate mass spectrometry. X-ray structures were obtained when possible.

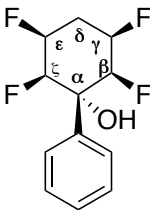
### 5.3.2 Fluorometabolite production in *C. elegans* in all-*cis* fluorocyclohexyl benzene derivatives

#### 5.3.2.1 All-*cis* (2,3,5,6-tetrafluorocyclohexyl)benzene (**62**)



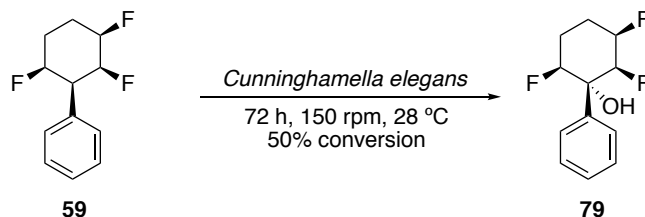
All-*cis* (2,3,5,6-tetrafluorocyclohexyl)benzene (**62**) was dissolved in DMF (50  $\mu$ L) and added to a mature culture of *C. elegans*. The compound was left to incubate with the fungus for 72 h. at 28 °C and 180 rpm. After 72 h, the fungus was centrifuged down, and the supernatant was extracted with EtOAc (3 x 50 mL). The combined organic phases were dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and evaporated under reduced pressure. Purification of the metabolite and residual starting material was achieved by reverse-phase HPLC, in a semi-preparative Kingsorb, and eluting with a mixture of 70:30 AcCN:water, both supplemented with 0.05% of TFA, at a flow rate of 1 mL/min. This afforded **78** at a  $t_{\text{ret}}$  = 18 min, and residual starting material **62** at  $t_{\text{ret}}$  = 23 min.

#### (1*S*, 2*R*, 3*S*, 5*R*, 6*S*)-2,3,5,6-tetrafluoro-1-phenylcyclohexan-1-ol (**78**)



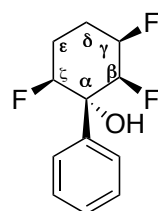
**<sup>1</sup>H NMR** (500 MHz, chloroform-*d*)  $\delta_{\text{H}}$  7.71 (2H, dd,  $J$  = 7.5, 1.8 Hz, *H*-Ar), 7.47 (3H, m, *H*-Ar), 5.10 (4H, m, *H*- $\beta$ , *H*- $\gamma$ , *H*- $\epsilon$ , *H*- $\zeta$ ), 2.87 (1H, m, *H*- $\delta_{\text{eq}}$ ), 2.49 (1H, m, *H*- $\delta_{\text{ax}}$ ), 2.01 (1H, s, -OH) **<sup>19</sup>F{<sup>1</sup>H} NMR** (470 MHz, chloroform-*d*)  $\delta_{\text{F}}$  -207.7 (2F, dd,  $J$  = 7.6, 5.1 Hz), -198.8 (2F, dd,  $J$  = 7.6, 5.0 Hz) **<sup>13</sup>C NMR** (126 MHz, chloroform-*d*)  $\delta_{\text{C}}$  138.7 (*C*-Ar, visible in HMBC), 129.7 (*C*-Ar), 129.3 (*C*-Ar), 126.1 (*C*-Ar), 90.5-85.2 (*C*- $\beta$ , *C*- $\gamma$ , *C*- $\epsilon$ , *C*- $\zeta$ ), 73.2 (*C*- $\alpha$ , visible in HMBC), 26.8 (*C*- $\delta$ ) **HRMS** (ESI<sup>-</sup>)  $m/z$  calc. for C<sub>12</sub>H<sub>11</sub>OF<sub>4</sub> [M-H]<sup>-</sup> 247.0746, found 247.0750.

### 5.3.2.2 All-*cis* (2,3,6-trifluorocyclohexyl)benzene (**59**)



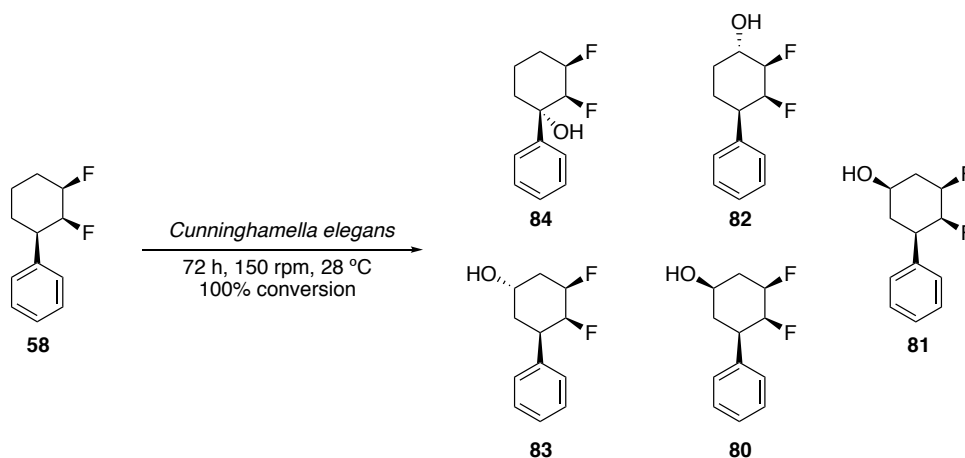
All-*cis* (2,3,6-trifluorocyclohexyl)benzene (**59**) was dissolved in DMF (50  $\mu$ L) and added to a mature culture of *C. elegans*. The compound was left to incubate with the fungus for 72 h. at 28 °C and 180 rpm. After 72 h, the fungus was centrifuged down, and the supernatant was extracted with EtOAc (3 x 50 mL). The combined organic phases were dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and evaporated under reduced pressure. Purification of the metabolite and residual starting material was achieved by reverse-phase HPLC, in a semi-preparative Kingsorb, and eluting with a mixture of 70:30 AcCN:water, both supplemented with 0.05% of TFA, at a flow rate of 1 mL/min. This afforded **79** at a  $t_{\text{ret}}$  = 19 min, and residual starting material **59** at  $t_{\text{ret}}$  = 24 min.

### (1*R*, 2*S*, 3*R*, 6*S*)-2,3,6-trifluoro-1-phenylcyclohexan-1-ol (**79**)



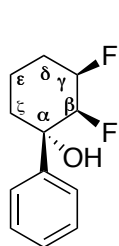
**<sup>1</sup>H NMR** (500 MHz, chloroform-*d*)  $\delta_{\text{H}}$  7.72 (2H, dd,  $J$  = 7.8, 1.6 Hz, *H*-Ar), 7.44 (3H, m, *H*-Ar), 5.15 (2H, m, *H*- $\gamma$ , *H*- $\zeta$ ), 4.90 (1H, d,  $J$  = 45.6 Hz, *H*- $\beta$ ), 2.32 (1H, m, *H*- $\delta_{\text{ax}}$ ), 2.22 (1H, m, *H*- $\epsilon_{\text{ax}}$ ), 2.11 (1H, m, *H*- $\epsilon_{\text{eq}}$ ), 1.96 (1H, m, *H*- $\delta_{\text{eq}}$ ), 1.87 (1H, s, -OH) **<sup>19</sup>F{<sup>1</sup>H} NMR** (470 MHz, chloroform-*d*)  $\delta_{\text{F}}$  -207.7 (*F*, dd,  $J$  = 24.3, 13.5 Hz), -191.6 (*F*, d,  $J$  = 13.5 Hz), -189.1 (*F*, d,  $J$  = 24.3 Hz) **<sup>13</sup>C NMR** (126 MHz, chloroform-*d*)  $\delta_{\text{C}}$  140.8(*C*-Ar), 129.1 (*C*-Ar), 126.2 (*C*-Ar), 90.8 (*C*- $\zeta$ , dd,  $J$  = 200.7, 12.2 Hz), 90.5 (*C*- $\beta$ , d,  $J$  = 183.3 Hz), 88.6 (*C*- $\gamma$ , dd,  $J$  = 200.7, 21.6 Hz), 73.8 (*C*- $\alpha$ , visible in HMBC), 24.7 (*C*- $\epsilon$ , dd,  $J$  = 21.6, 12.2 Hz), 19.9 (*C*- $\delta$ , d,  $J$  = 20.5 Hz) **HRMS** (ESI<sup>+</sup>)  $m/z$  calc. for C<sub>12</sub>H<sub>13</sub>OF<sub>3</sub>Na<sup>+</sup> [M+Na]<sup>+</sup> 253.0816, found 253.0807.

### 5.3.2.3 All-*cis* (2,3-difluorocyclohexyl)benzene (**58**)



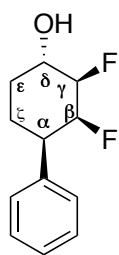
All-*cis* (2,3-difluorocyclohexyl)benzene (**58**) was dissolved in DMF (50  $\mu$ L) and added to a mature culture of *C. elegans*. The compound was left to incubate with the fungus for 72 h. at 28  $^{\circ}$ C and 180 rpm. After 72 h, the fungus was centrifuged down, and the supernatant was extracted with EtOAc (3 x 50 mL). The combined organic phases were dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and evaporated under reduced pressure. Purification of the metabolite and residual starting material was achieved by reverse-phase HPLC, in a semi-preparative Kingsorb, and eluting with a mixture of 70:30 AcCN:water, both supplemented with 0.05% of TFA, at a flow rate of 1 mL/min. This afforded **80** at a  $t_{\text{ret}}$  = 16 min, **81** at a  $t_{\text{ret}}$  = 16.5 min, **82** at a  $t_{\text{ret}}$  = 17.5 min, **83** at a  $t_{\text{ret}}$  = 18 min and **84** at a  $t_{\text{ret}}$  = 22 min.

#### (1*S*, 2*S*, 3*R*)-2,3-difluoro-1-phenylcyclohexan-1-ol (**84**)



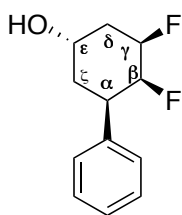
**<sup>1</sup>H NMR** (500 MHz, chloroform-*d*)  $\delta_{\text{H}}$  7.56 (2H, m, H-Ar), 7.41 (2H, m, H-Ar), 7.35 (1H, m, H-Ar), 5.05 (1H, m, H- $\gamma$ ) 4.78 (1H, dd,  $J$  = 50.7, 8.9 Hz, H- $\beta$ ), 2.26 (1H, m, H- $\zeta$ ), 2.03 (1H, m, H- $\delta_{\text{eq}}$ ), 1.89 (1H, m, H- $\delta_{\text{ax}}$ ), 1.82 (2H, m, H- $\epsilon$ ), 1.75 (1H, d,  $J$  = 15.1 Hz, H- $\zeta$ ) **<sup>19</sup>F{<sup>1</sup>H} NMR** (470 MHz, chloroform-*d*)  $\delta_{\text{F}}$  -205.3 (F, d,  $J$  = 16.4 Hz), -189.8 (F, d,  $J$  = 16.3 Hz) **<sup>13</sup>C NMR** (126 MHz, chloroform-*d*)  $\delta_{\text{C}}$  144.2 (C-Ar, visible in HMBC), 129.0 (C-Ar), 128.2 (C-Ar), 125.5 (C-Ar), 91.7 (C- $\beta$ , dd,  $J$  = 198.4, 15.9 Hz), 89.7 (C- $\gamma$ , dd,  $J$  = 196.2, 18.1 Hz), 74.7 (C- $\alpha$ , visible in HMBC), 31.2 (C- $\zeta$ ), 25.1 (C- $\delta$ , dd,  $J$  = 21.9, 3.1 Hz), 17.9 (C- $\epsilon$ , d,  $J$  = 11.9 Hz) **HRMS** (ESI+)  $m/z$  calc. for C<sub>12</sub>H<sub>18</sub>ONF<sub>2</sub> [M+NH<sub>4</sub>]<sup>+</sup> 230.1351, found 230.1355.

**(1*S*, 2*R*, 3*S*, 4*R*)-2,3-difluoro-4-phenylcyclohexan-1-ol (82)**



**<sup>1</sup>H NMR** (500 MHz, chloroform-*d*)  $\delta_{\text{H}}$  7.35 (5H, m, *H*-Ar), 5.04 (1H, dd,  $J = 60.8, 9.1$  Hz, *H*- $\beta$ ) 4.41 (1H, dddd,  $J = 47.7, 26.8, 9.3, 2.3$  Hz, *H*- $\gamma$ ), 4.25 (1H, m, *H*- $\delta$ ), 2.77 (1H, ddd,  $J = 35.7, 13.2, 4.2$  Hz, *H*- $\alpha$ ), 2.22 (1H, m, *H*- $\epsilon_{\text{ax}}$ ), 2.10 (1H, m, *H*- $\zeta_{\text{ax}}$ ), 1.82 (1H, dddd,  $J = 14.0, 6.1, 3.0, 1.6$  Hz, *H*- $\zeta_{\text{eq}}$ ) 1.54 (1H, qd,  $J = 13.2, 4.2$  Hz, *H*- $\epsilon_{\text{eq}}$ ) **<sup>19</sup>F{<sup>1</sup>H} NMR** (470 MHz, chloroform-*d*)  $\delta_{\text{F}}$  -210.6 (*F*, d,  $J = 17.0$  Hz), -195.36 (*F*, d,  $J = 17.4$  Hz) **<sup>13</sup>C NMR** (126 MHz, chloroform-*d*)  $\delta_{\text{C}}$  140.2 (*C*-Ar, visible in HMBC), 128.6 (*C*-Ar), 127.7 (*C*-Ar), 127.3 (*C*-Ar), 95.8 (*C*- $\gamma$ , dd,  $J = 184.6, 18.2$  Hz), 92.4 (*C*- $\beta$ , dd,  $J = 181.8, 16.3$  Hz), 68.7 (*C*- $\delta$ , dd,  $J = 18.8, 3.0$  Hz), 45.68 (*C*- $\alpha$ , dd,  $J = 19.2, 5.2$  Hz), 30.45 (*C*- $\epsilon$ , d,  $J = 7.7$  Hz), 24.24 (*C*- $\zeta$ , d,  $J = 4.4$  Hz) **HRMS** (ESI<sup>+</sup>)  $m/z$  calc. for C<sub>12</sub>H<sub>14</sub>OF<sub>2</sub>Na [M+Na]<sup>+</sup> 235.0910, found 235.0903.

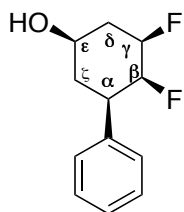
**(1*S*, 3*R*, 4*S*, 5*R*)-3,4-difluoro-5-phenylcyclohexan-1-ol (83)**



**<sup>1</sup>H NMR** (500 MHz, chloroform-*d*)  $\delta_{\text{H}}$  7.32 (5H, m, *H*-Ar), 5.00 (2H, m, *H*- $\beta$ , *H*- $\gamma$ ) 4.46 (1H, s, *H*- $\epsilon$ ), 3.25 (1H, ddd,  $J = 36.2, 13.4, 4.2$  Hz, *H*- $\alpha$ ), 2.19 (3H, m, *H*- $\delta$ , *H*- $\zeta_{\text{ax}}$ ), 1.82 (1H, d,  $J = 14.2$  Hz, *H*- $\zeta_{\text{eq}}$ ) **<sup>19</sup>F{<sup>1</sup>H} NMR** (470 MHz, chloroform-*d*)  $\delta_{\text{F}}$  -217.5 (*F*, d,  $J = 16.0$  Hz), -193.9 (*F*, d,  $J = 16.0$  Hz) **<sup>13</sup>C NMR** (126 MHz, chloroform-*d*)  $\delta_{\text{C}}$  140.3 (*C*-Ar, visible in HMBC), 128.6 (*C*-Ar), 128.1 (*C*-Ar), 127.2 (*C*-Ar), 92.0 (*C*- $\beta$ , dd,  $J = 199.6, 17.1$  Hz), 88.3 (*C*- $\gamma$ , dd,  $J = 199.6, 19.4$  Hz), 66.2 (*C*- $\epsilon$ , d,  $J = 15.3$  Hz), 39.6 (*C*- $\alpha$ , dd,  $J = 24.7, 5.8$  Hz), 32.6 (*C*- $\delta$ , *C*- $\zeta$ , m) **HRMS** (ESI<sup>+</sup>)  $m/z$  calc. for C<sub>12</sub>H<sub>18</sub>ONF<sub>2</sub> [M+NH<sub>4</sub>]<sup>+</sup> 230.1351, found 230.1356.



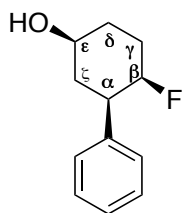
**(1*R*, 3*R*, 4*S*, 5*R*)-3,4-difluoro-5-phenylcyclohexan-1-ol (80)**



**<sup>1</sup>H NMR** (500 MHz, chloroform-*d*)  $\delta_{\text{H}}$  7.36 (2H, m, *H*-Ar), 7.30 (3H, m, *H*-Ar), 4.91 (1H, dd,  $J = 52.6, 8.4$  Hz, *H*- $\beta$ ), 4.61 (1H, m, *H*- $\gamma$ ), 3.89 (1H, m, *H*- $\epsilon$ ), 2.66 (1H, m, *H*- $\alpha$ ), 2.40 (1H, dq,  $J = 10.1, 4.9$  Hz, *H*- $\delta_{\text{eq}}$ ), 2.04 (3H, m, *H*- $\delta_{\text{ax}}$ , *H*- $\zeta$ ) **<sup>19</sup>F{<sup>1</sup>H} NMR** (470 MHz, chloroform-*d*)  $\delta_{\text{F}}$  -213.6 (*F*, d,  $J = 16.1$  Hz), -186.8 (*F*, d,  $J = 16.1$  Hz)

**<sup>13</sup>C NMR** (126 MHz, chloroform-*d*)  $\delta_{\text{C}}$  139.8 (*C*-Ar), 128.7 (*C*-Ar), 127.9 (*C*-Ar), 127.3 (*C*-Ar), 90.6 (*C*- $\beta$ , dd,  $J = 160.8, 16.8$  Hz), 89.1 (*C*- $\gamma$ , dd,  $J = 160.8, 19.2$  Hz), 67.1 (*C*- $\epsilon$ , d,  $J = 15.3$  Hz), 41.7 (*C*- $\alpha$ , dd,  $J = 25.9, 6.8$  Hz), 35.1 (*C*- $\delta$ , dd,  $J = 21.6, 2.9$  Hz), 34.5 (*C*- $\zeta$ , d,  $J = 3.5$  Hz) **HRMS** (ESI<sup>+</sup>)  $m/z$  calc. for C<sub>12</sub>H<sub>18</sub>ONF<sub>2</sub> [M+NH<sub>4</sub>]<sup>+</sup> 230.1351, found 230.1356.

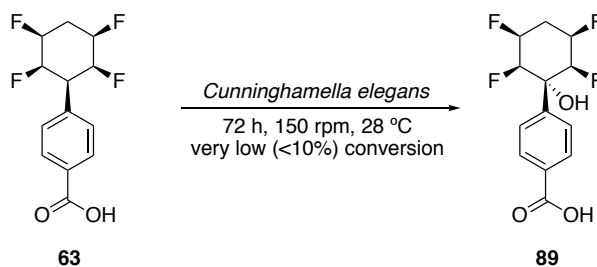
**(1*S*, 3*R*, 4*R*)-4-fluoro-3-phenylcyclohexan-1-ol (81)**



**<sup>1</sup>H NMR** (500 MHz, chloroform-*d*)  $\delta_{\text{H}}$  7.33 (5H, m, *H*-Ar), 4.79 (1H, ddd,  $J = 49.0, 3.8, 1.9$  Hz, *H*- $\beta$ ), 3.85 (1H, m, *H*- $\epsilon$ ), 2.78 (1H, m, *H*- $\alpha$ ), 2.24 (1H, m, *H*- $\gamma_{\text{ax}}$ ), 2.07 (2H, m, *H*- $\zeta$ ), 1.94 (1H, m, *H*- $\delta_{\text{eq}}$ ), 1.70 (2H, m, *H*- $\delta_{\text{ax}}$ , *H*- $\gamma_{\text{eq}}$ ) **<sup>19</sup>F{<sup>1</sup>H} NMR** (470 MHz, chloroform-*d*)  $\delta_{\text{F}}$  -197.5 (*F*, s) **<sup>13</sup>C NMR** (126 MHz, chloroform-*d*)  $\delta_{\text{C}}$  141.5 (*C*-Ar, visible in

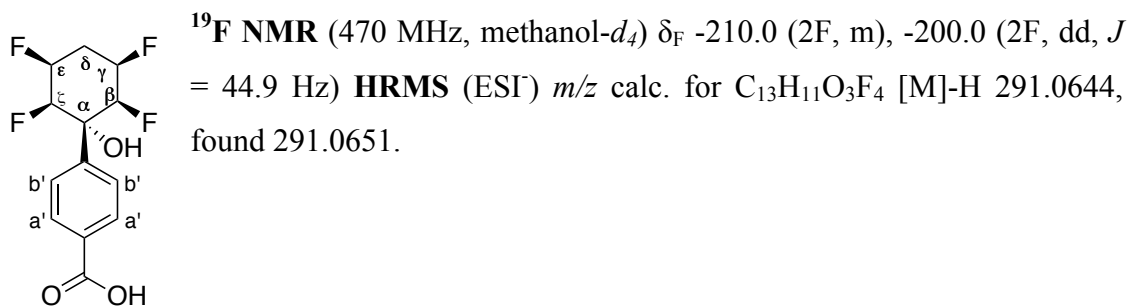
HMBC), 128.5 (*C*-Ar), 128.0 (*C*-Ar), 126.8 (*C*-Ar), 90.4 (*C*- $\beta$ , d,  $J = 175.5$  Hz), 70.1 (*C*- $\epsilon$ , s), 45.9 (*C*- $\alpha$ , d,  $J = 19.9$  Hz), 35.5 (*C*- $\zeta$ , d,  $J = 2.9$  Hz), 29.7 (*C*- $\gamma$ , d,  $J = 21.4$  Hz), 19.1 (*C*- $\delta$ , s) **HRMS** (ESI<sup>+</sup>)  $m/z$  calc. for C<sub>12</sub>H<sub>19</sub>ONF [M+NH<sub>4</sub>]<sup>+</sup> 212.1445, found 212.1445.

#### 5.3.2.4 All-*cis* 4-(2,3,5,6-tetrafluorocyclohexyl)benzoic acid (**63**)

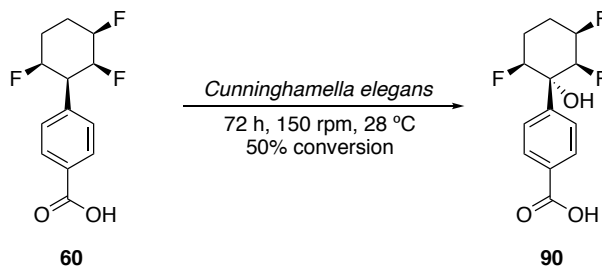


All-*cis* 4-(2,3,5,6-tetrafluorocyclohexyl)benzoic acid (**63**) was dissolved in DMF (50  $\mu$ L) and added to a mature culture of *C. elegans*. The compound was left to incubate with the fungus for 72 h. at 28  $^{\circ}$ C and 180 rpm. After 72 h, the fungus was centrifuged down, and the supernatant was extracted with EtOAc (3 x 50 mL). The combined organic phases were dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and evaporated under reduced pressure. Purification of the metabolite and residual starting material was achieved by reverse-phase HPLC, in a semi-preparative Kingsorb, and eluting with a mixture of 50:50 AcCN:water, both supplemented with 0.05% of TFA, at a flow rate of 1 mL/min. This afforded **89** at a  $t_{\text{ret}}$  = 24 min, and starting material **63** at a  $t_{\text{ret}}$  = 28 min.

#### 4-((1*S*, 2*R*, 3*S*, 5*R*, 6*S*)-2,3,5,6-tetrafluoro-1-hydroxycyclohexyl)benzoic acid (**89**)

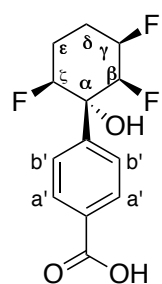


#### 5.3.2.5 All-*cis* 4-(2,3,6-trifluorocyclohexyl)benzoic acid (**60**)



All-*cis* 4-(2,3,6-trifluorocyclohexyl)benzoic acid (**60**) was dissolved in DMF (50  $\mu$ L) and added to a mature culture of *C. elegans*. The compound was left to incubate with the fungus for 72 h. at 28 °C and 180 rpm. After 72 h, the fungus was centrifuged down, and the supernatant was extracted with EtOAc (3 x 50 mL). The combined organic phases were dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and evaporated under reduced pressure. Purification of the metabolite and residual starting material was achieved by reverse-phase HPLC, in a semi-preparative Kingsorb, and eluting with a mixture of 50:50 AcCN:water, both supplemented with 0.05% of TFA, at a flow rate of 1 mL/min. This afforded **90** at a  $t_{\text{ret}}$  = 23 min, and starting material **60** at a  $t_{\text{ret}}$  = 27.5 min.

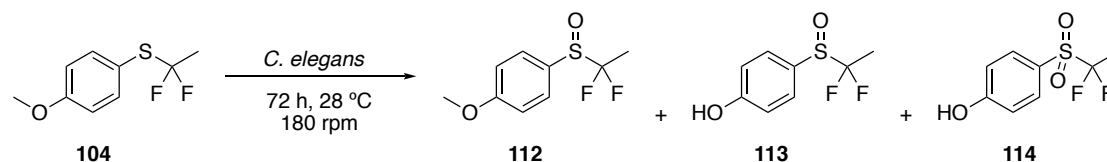
#### 4-((1R, 2S, 3R, 6S)-2,3,6-trifluoro-1-hydroxycyclohexyl)benzoic acid (**90**)



**<sup>1</sup>H NMR** (500 MHz, methanol-*d*<sub>4</sub>)  $\delta_{\text{H}}$  8.06 (2H, m, *H*-Ar-*a*'), 7.85 (2H, d, *J* = 8.5 Hz, *H*-Ar-*b*'), 5.11 (2H, m, *H*- $\gamma$ , *H*- $\zeta$ ), 4.79 (1H, d, *J* = 46.9 Hz, *H*- $\beta$ ), 2.21 (3H, m, *H*- $\epsilon$ , *H*- $\delta_{\text{eq}}$ ), 1.91 (1H, m,  $\delta_{\text{ax}}$ ) **<sup>19</sup>F{<sup>1</sup>H} NMR** (470 MHz, methanol-*d*<sub>4</sub>)  $\delta_{\text{F}}$  -209.9 (*F*, dd, *J* = 23.0, 12.8 Hz), -192.9 (*F*, d, *J* = 12.8 Hz), -191.2 (*F*, d, *J* = 23.0 Hz) **<sup>13</sup>C NMR** (126 MHz, methanol-*d*<sub>4</sub>)  $\delta_{\text{C}}$  169.4 (COOH, visible in HMBC), 149.1 (*C*-Ar, visible in HMBC), 132.0 (*C*-Ar, visible in HMBC), 129.0 (*C*-Ar), 126.6 (*C*-Ar), 91.2 (*C*- $\zeta$ , dd, *J* = 203.6, 15.5 Hz), 90.8 (*C*- $\beta$ , d, *J* = 178.3 Hz), 88.5 (*C*- $\gamma$ , dd, *J* = 198.2, 19.9 Hz), 73.0 (*C*- $\alpha$ , visible in HMBC), 24.5 (*C*- $\epsilon$ , dd, *J* = 34.4, 12.6 Hz), 19.7 (*C*- $\delta$ , dd, *J* = 24.1, 4.1 Hz) **HRMS** (ESI) *m/z* calc. for C<sub>13</sub>H<sub>12</sub>O<sub>3</sub>F<sub>3</sub> [M]-H 273.0739, found 273.0743.

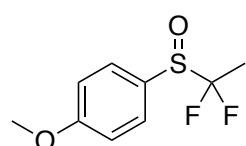
### 5.3.3 Fluorometabolite production in *C. elegans* in $\alpha,\alpha$ -difluoroethyl thio- and oxy- ethers

#### 5.3.3.1 (1,1-Difluoroethyl)(4-methoxyphenyl)sulfane (**104**)



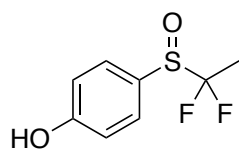
(1,1-Difluoroethyl)(4-methoxyphenyl)sulfane (**104**) was dissolved in DMF (50  $\mu$ L) and added to a mature culture of *C. elegans*. The compound was left to incubate with the fungus for 72 h. at 28  $^\circ$ C and 180 rpm. After 72 h, the fungus was centrifuged down, and the supernatant was extracted with DCM (3 x 50 mL) and Et<sub>2</sub>O (3 x 50 mL). The combined organic phases were dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and evaporated under reduced pressure. Purification of the metabolite and residual starting material was achieved by reverse-phase HPLC, in a Phenomenex Luna, and eluting with a mixture of 60:40 AcCN:water, both supplemented with 0.05% of TFA, at a flow rate of 1 mL/min. This afforded **112** at a  $t_{\text{ret}}$  = 24 min, **113** at a  $t_{\text{ret}}$  = 16 min and **114** at a  $t_{\text{ret}}$  = 21 min.

#### 1-((1,1-Difluoroethyl)sulfinyl)-4-methoxybenzene (**112**)



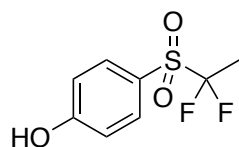
**<sup>1</sup>H NMR** (500 MHz, chloroform-*d*)  $\delta_{\text{H}}$  7.64 (2H, d,  $J$  = 8.9 Hz,  $H$ -Ar), 7.06 (2H, d,  $J$  = 8.9 Hz,  $H$ -Ar), 3.88 (3H, s, OCH<sub>3</sub>), 1.75 (3H, t,  $J$  = 18.4 Hz, CF<sub>2</sub>CH<sub>3</sub>); **<sup>19</sup>F{<sup>1</sup>H} NMR** (470 MHz, chloroform-*d*)  $\delta_{\text{F}}$  -94.2 (d,  $J$  = 225.8 Hz), - 97.6 (d,  $J$  = 225.8 Hz); **<sup>13</sup>C NMR** (126 MHz, chloroform-*d*)  $\delta_{\text{C}}$  163.1 (s, C-4), 128.1 (t,  $J$  = 323.8 Hz, CF<sub>2</sub>), 127.6 (s, C-2), 114.8 (s, C-3), 55.6 (s, OCH<sub>3</sub>), 16.5 (t,  $J$  = 22.3 Hz, CF<sub>2</sub>CH<sub>3</sub>); **HRMS** (ESI<sup>+</sup>)  $m/z$  calc. for C<sub>9</sub>H<sub>11</sub>O<sub>2</sub>F<sub>2</sub>S [M+H]<sup>+</sup> 221.0448, found 221.0440.

#### 4-((1,1-Difluoroethyl)sulfinyl)phenol (113)



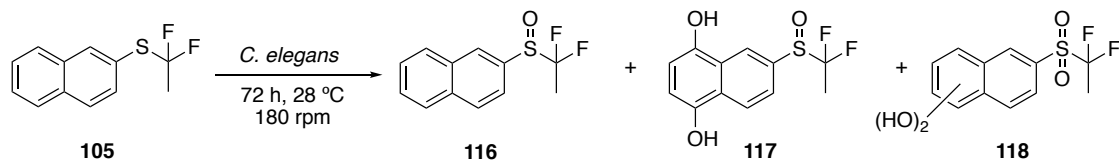
$^1\text{H}$  NMR (500 MHz, chloroform-*d*)  $\delta_{\text{H}}$  7.60 (2H, d,  $J = 8.9$  Hz, *H*-Ar), 7.01 (2H, d,  $J = 8.9$  Hz, *H*-Ar), 1.75 (3H, t,  $J = 18.4$  Hz,  $\text{CF}_2\text{CH}_3$ );  $^{19}\text{F}\{^1\text{H}\}$  NMR (470 MHz, chloroform-*d*)  $\delta_{\text{F}}$  -94.1 (d,  $J = 225.1$  Hz), -97.6 (d,  $J = 225.1$  Hz);  $^{13}\text{C}$  NMR (126 MHz, chloroform-*d*)  $\delta_{\text{C}}$  159.7 (s, C-4), 127.9 (s, C-2), 125.4 (t,  $J = 276.2$  Hz,  $\text{CF}_2$ , visible in HMBC), 116.4 (s, C-3), 16.5 (t,  $J = 22.3$  Hz,  $\text{CF}_2\text{CH}_3$ ); HRMS (ESI $^+$ )  $m/z$  calc. for  $\text{C}_8\text{H}_9\text{O}_2\text{F}_2\text{S}$   $[\text{M}+\text{H}]^+$  207.0291, found 207.0286.

#### 4-((1,1-Difluoroethyl)sulfonyl)phenol (114)



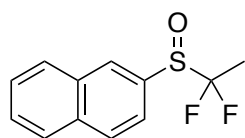
$^1\text{H}$  NMR (500 MHz, chloroform-*d*)  $\delta_{\text{H}}$  7.87 (2H, d,  $J = 8.9$  Hz, *H*-Ar), 7.01 (2H, d,  $J = 8.9$  Hz, *H*-Ar), 2.02 (3H, t,  $J = 18.3$  Hz,  $\text{CF}_2\text{CH}_3$ );  $^{19}\text{F}\{^1\text{H}\}$  NMR (470 MHz, chloroform-*d*)  $\delta_{\text{F}}$  -97.2 (s);  $^{13}\text{C}$  NMR (126 MHz, chloroform-*d*)  $\delta_{\text{C}}$  161.8 (s, C-4), 133.5 (s, C-3), 123.6 (t,  $J = 262.9$  Hz,  $\text{CF}_2$ , visible in HMBC), 116.3 (s, C-2), 16.6 (t,  $J = 22.4$  Hz,  $\text{CF}_2\text{CH}_3$ ); HRMS (ESI $^+$ )  $m/z$  calc. for  $\text{C}_8\text{H}_8\text{O}_3\text{F}_2\text{SNa}$   $[\text{M}+\text{Na}]^+$  245.0060, found 245.0055; HRMS (ESI $^-$ )  $m/z$  calc. for  $\text{C}_8\text{H}_7\text{O}_3\text{F}_2\text{S}$   $[\text{M}-\text{H}]^-$  221.0084, found 221.0087.

### 5.3.3.2 (1,1-Difluoroethyl)(naphthalene-2-yl) sulfane (105)



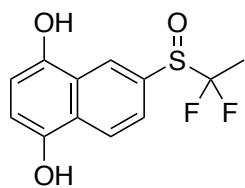
(1,1-Difluoroethyl)(naphthalene-2-yl) sulfane (**105**) was dissolved in DMF (50  $\mu\text{L}$ ) and added to a mature culture of *C. elegans*. The compound was left to incubate with the fungus for 72 h. at 28  $^\circ\text{C}$  and 180 rpm. After 72 h, the fungus was centrifuged down, and the supernatant was extracted with DCM (3 x 50 mL) and Et<sub>2</sub>O (3 x 50 mL). The combined organic phases were dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and evaporated under reduced pressure. Purification of the metabolite and residual starting material was achieved by reverse-phase HPLC, in a Phenomenex Luna, and eluting with a mixture of 60:40 AcCN:water, both supplemented with 0.05% of TFA, at a flow rate of 1 mL/min. This afforded **116** at a  $t_{\text{ret}} = 37$  min, **117** at a  $t_{\text{ret}} = 13$  min and **118** at a  $t_{\text{ret}} = 17$  min.

### 2-((1,1-Difluoroethyl)sulfinyl)naphthalene (**116**)



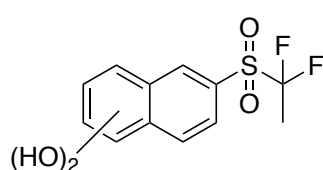
**<sup>1</sup>H NMR** (500 MHz, chloroform-*d*):  $\delta_{\text{H}}$  8.28 (s, 1H), 8.01 (d,  $J = 8.7$  Hz, 1H), 7.99 – 7.92 (m, 2H), 7.69 (ddt,  $J = 8.7, 2.6, 1.3$  Hz, 1H), 7.67 – 7.60 (m, 2H), 1.77 (t,  $J = 18.5$  Hz, 3H); **<sup>19</sup>F NMR** (471 MHz, chloroform-*d*):  $\delta_{\text{F}}$  -92.9 (d,  $J = 227.0$  Hz), -95.6 (d,  $J = 227.0$  Hz); **HRMS** (ESI<sup>+</sup>)  $m/z$  calc. for C<sub>12</sub>H<sub>11</sub>OF<sub>2</sub>S [M+H]<sup>+</sup> 241.0420, found 241.0491. Data is consistent with the full characterisation carried out for **116** from chemical reaction (**Section 5.2.9**).

**6-((1,1-Difluoroethyl)sulfinyl)naphthalene-1,4-diol (117)**



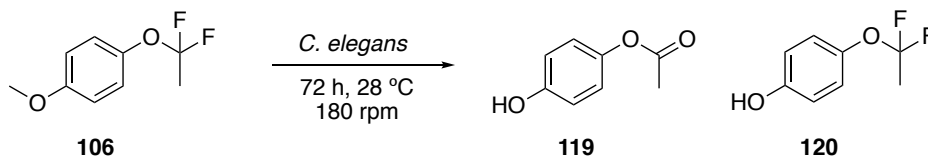
**<sup>1</sup>H NMR** (500 MHz, chloroform-*d*):  $\delta_{\text{H}}$  7.81 (1H, d,  $J = 7.9$  Hz, *H*-Ar), 7.61 (1H, d,  $J = 7.9$  Hz, *H*-Ar), 7.44 (1H, s, *H*-Ar), 6.50 (1H, d,  $J = 11.0$  Hz, *H*-Ar), 6.12 (1H, d,  $J = 11.0$  Hz, *H*-Ar), 1.80 (3H, t,  $J = 18.4$  Hz,  $\text{CF}_2\text{CH}_3$ ); **<sup>19</sup>F NMR** (471 MHz, chloroform-*d*):  $\delta_{\text{F}}$  -93.2(d,  $J = 226.9$  Hz), -96.7 (d,  $J = 226.9$  Hz); **HRMS** (ESI<sup>-</sup>)  $m/z$  calc. for  $\text{C}_{12}\text{H}_9\text{O}_3\text{F}_2\text{S}$   $[\text{M}-\text{H}]^-$  271.0246, found 271.0245.

**General (1,1-difluoroethyl)sulfinyl)naphthalene (118)**



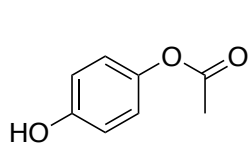
**<sup>19</sup>F NMR** (471 MHz, chloroform-*d*):  $\delta_{\text{F}}$  -96.8 (s);  $m/z$  calc. for  $\text{C}_{12}\text{H}_{10}\text{O}_4\text{F}_2\text{S}$   $[\text{M}]^+$  288, found 288.


#### 5.3.3.3 1-(1,1-Difluoroethoxy)-4-methoxybenzene (106)



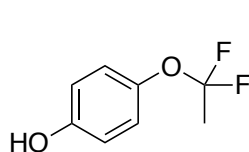
1-(1,1-Difluoroethoxy)-4-methoxybenzene (**106**) was dissolved in DMF (50  $\mu$ L) and added to a mature culture of *C. elegans*. The compound was left to incubate with the fungus for 72 h. at 28 °C and 180 rpm. After 72 h, the fungus was centrifuged down, and the supernatant was extracted with EtOAc (3 x 50 mL). The combined organic phases were dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and evaporated under reduced pressure. Purification of the metabolite and residual starting material was achieved by reverse-phase HPLC, in a Phenomenex Luna, and eluting with a mixture of 60:40 AcCN:water, both supplemented with 0.05% of TFA, at a flow rate of 1 mL/min. This afforded **119** at a  $t_{\text{ret}}$  = 25 min, and **120** at a  $t_{\text{ret}}$  = 19 min.


### 4-Acetoxyphenol (119)



 **<sup>1</sup>H NMR** (500 MHz, chloroform-*d*)  $\delta_{\text{H}}$  6.95 (2H, d,  $J = 8.9$  Hz, *H*-Ar), 6.82 (2H, d,  $J = 8.9$  Hz, *H*-Ar), 2.28 (3H, t,  $J = 18.4$  Hz, COCH<sub>3</sub>); **<sup>13</sup>C NMR** (126 MHz, chloroform-*d*)  $\delta_{\text{C}}$  160.8 (s, CO, visible in HMBC), 122.7 (s, *C*-Ar), 116.1 (s, *C*-Ar), 21.1 (s, COCH<sub>3</sub>). Data is in good agreement with the literature values.<sup>9</sup>

### 4-(1,1-Difluoro)ethoxy phenol (120)

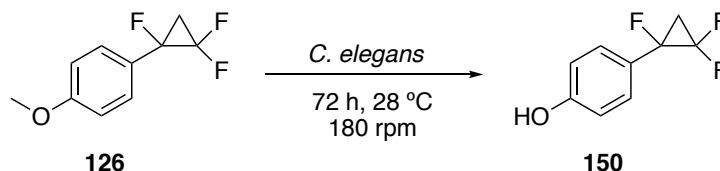


 **<sup>1</sup>H NMR** (500 MHz, chloroform-*d*)  $\delta_{\text{H}}$  6.75 (2H, d,  $J = 2.7$  Hz, *H*-Ar), 6.64 (2H, d,  $J = 2.7$  Hz, *H*-Ar), 1.87 (3H, t,  $J = 13.3$  Hz,  $\text{CF}_2\text{CH}_3$ ); **<sup>19</sup>F{<sup>1</sup>H} NMR** (470 MHz, chloroform-*d*)  $\delta_{\text{F}}$  -64.6 (s); **<sup>13</sup>C NMR** (126 MHz, chloroform-*d*)  $\delta_{\text{C}}$  114.4 (s, *C*-Ar), 110.1 (s, *C*-Ar), 22.5 (t,  $J = 32.0$  Hz,  $\text{CF}_2\text{CH}_3$ ). Low amounts did not allow a full characterisation.



### 5.3.4 Fluorometabolite production in *C. elegans* in $\alpha,\beta,\beta$ -(trifluorocyclopropyl) benzene derivatives

#### 5.3.4.1 1-Methoxy-4-(1,2,2-trifluorocyclopropyl)benzene (**126**)



1-Methoxy-4-(1,2,2-trifluorocyclopropyl)benzene (**126**) was dissolved in DMF (50  $\mu$ L) and added to a mature culture of *C. elegans*. The compound was left to incubate with the fungus for 72 h. at 28 °C and 180 rpm. After 72 h, the fungus was centrifuged down, and the supernatant was extracted with EtOAc (3 x 50 mL). The combined organic phases were dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and evaporated under reduced pressure. Purification of the metabolite and residual starting material was attempted by column chromatography, starting with 100% petroleum ether and following with 10% EtOAc, 20% EtOAc, 30% EtOAc and 50% EtOAc, finishing with 100% EtOAc wash. The purification allowed the isolation of **150**.

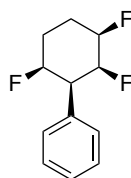
#### 4-(1,2,2-trifluorocyclopropyl)phenol (**150**)

**<sup>1</sup>H NMR** (500 MHz, chloroform-*d*)  $\delta_{\text{H}}$  7.36 (2H, dt,  $J = 8.7, 1.3$  Hz,  $H\text{-Ar}$ ), 6.88 (2H, dt,  $J = 8.7, 1.3$  Hz,  $H\text{-Ar}$ ), 2.14-1.93 (2H, m,  $\text{CH}_2$ ); **<sup>19</sup>F{<sup>1</sup>H} NMR** (470 MHz, chloroform-*d*)  $\delta_{\text{F}}$  -135.7 (dd,  $J = 166.9, 9.1$  Hz), -142.5 (d,  $J = 166.9$  Hz), -175.3 (s); **<sup>13</sup>C NMR** (126 MHz, chloroform-*d*)  $\delta_{\text{C}}$  157.4 (s,  $\text{C-Ar}$ , visible in HMBC), 129.6 (s,  $\text{C-Ar}$ ), 123.6 (s,  $\text{C-Ar}$ , visible in HMBC), 115.4 (s,  $\text{C-Ar}$ ), 32.7 (d,  $J = 175.0$  Hz, CF), 29.6-29.0 (m,  $\text{CF}_2$ ), 22.1 (m,  $\text{CH}_2$ ); **HRMS** (ESI<sup>-</sup>)  $m/z$  calc. for C<sub>9</sub>H<sub>6</sub>OF<sub>3</sub> [M-H]<sup>-</sup> 187.0376, found 187.0372.

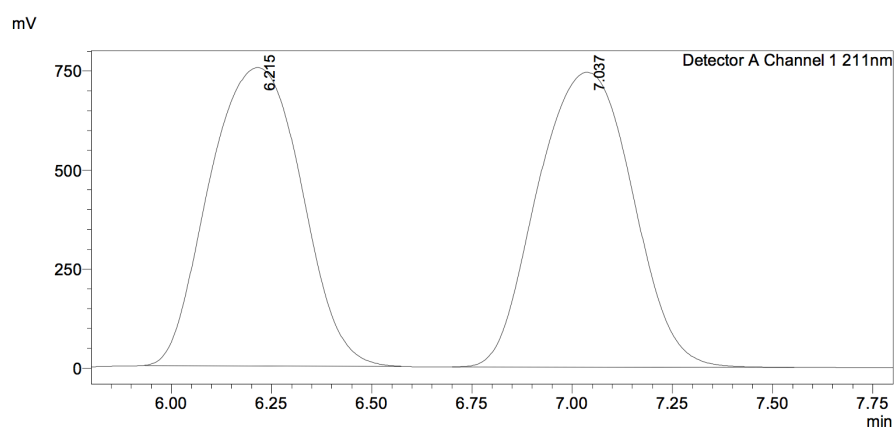
## 5.4 Enantiomeric excess analysis

### 5.4.1 Enantiomeric excess analysis in all-*cis* fluorocyclohexyl derivatives

#### 5.4.1.1 All-*cis* (2,3,6-trifluorocyclohexyl)benzene (**59**)

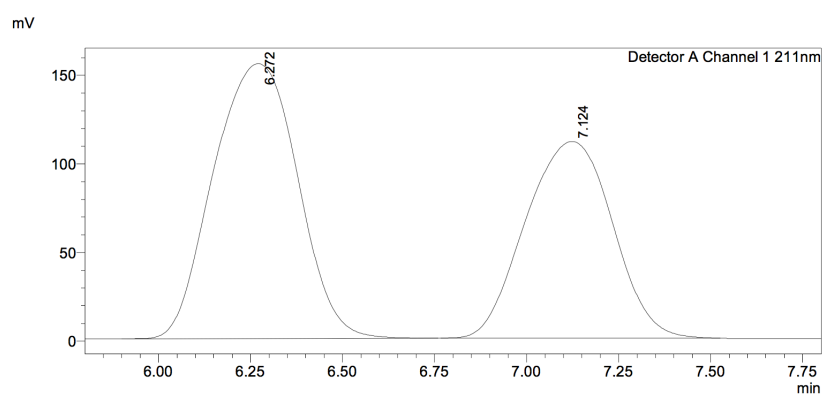


HPLC data for compound **59**: Chiralcel IC (95:5 hexane:IPA, flow rate 1 mLmin<sup>-1</sup>, 211 nm, 30 °C),  $t_R$  (A): 6.2 min,  $t_R$  (B): 7.0 min. 57:43 ee.



#### <Peak Table>

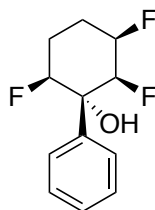
Detector A Channel 1 211nm		
Peak#	Ret. Time	Area%
1	6.215	49.549
2	7.037	50.451
Total		100.000



#### <Peak Table>

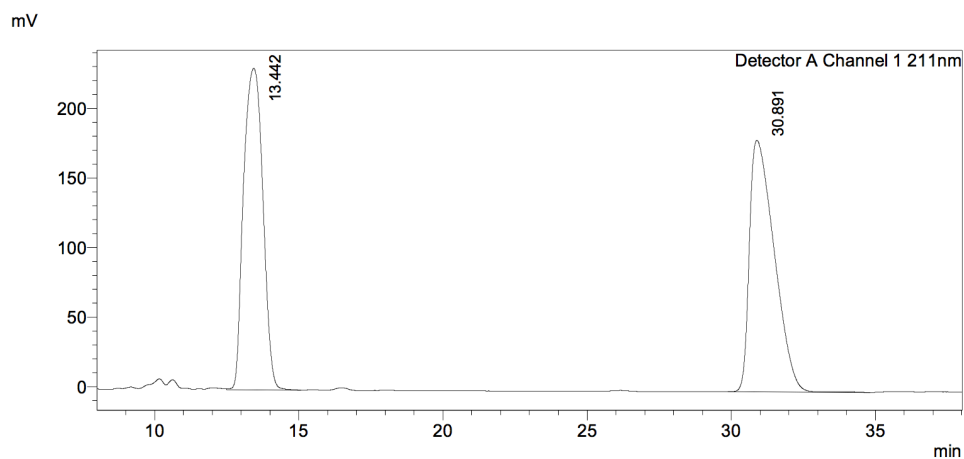
Detector A Channel 1 211nm		
Peak#	Ret. Time	Area%
1	6.272	57.487
2	7.124	42.513
Total		100.000

#### 5.4.1.2 2,3,6-Trifluoro-1-phenylcyclohexan-1-ol (79)



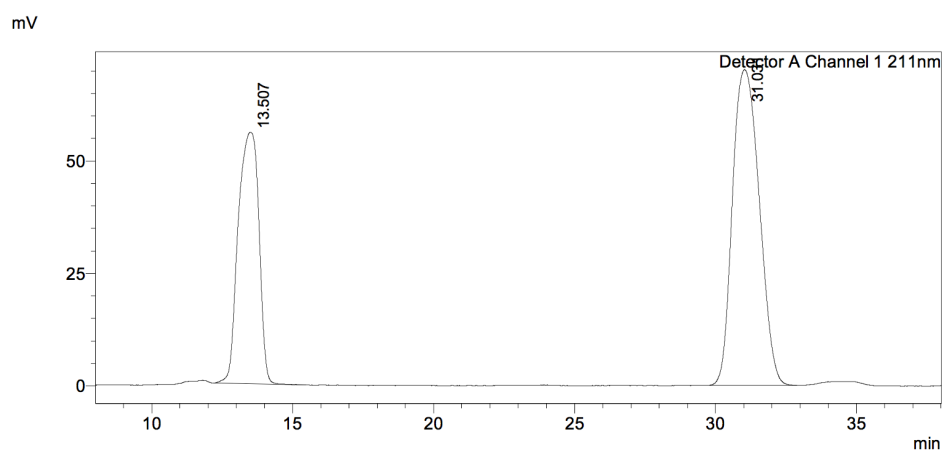
HPLC data for compound **79**: Chiralcel ID (95:5 hexane:IPA, flow rate 1 mLmin<sup>-1</sup>, 211 nm, 30 °C), *t<sub>R</sub>* (A): 13.5 min, *t<sub>R</sub>* (B): 31.0 min; 38:62 ee.

##### <Chromatogram>



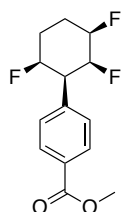
Detector A Channel 1 211nm		
Peak#	Ret. Time	Area%
1	13.442	49.911
2	30.891	50.089
Total		100.000

##### <Chromatogram>



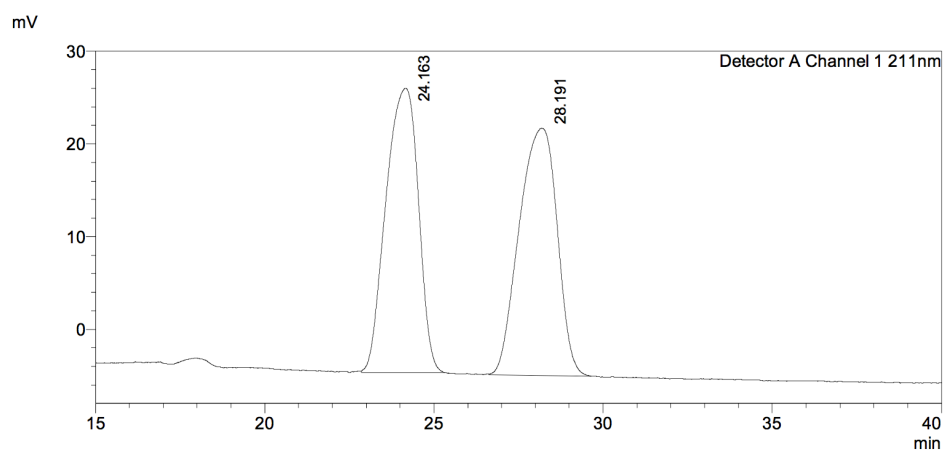
Detector A Channel 1 211nm		
Peak#	Ret. Time	Area%
1	13.507	38.116
2	31.031	61.884
Total		100.000

### 5.4.1.3 All-cis methyl 4-(2,3,6-trifluorocyclohexyl) benzoate (93)



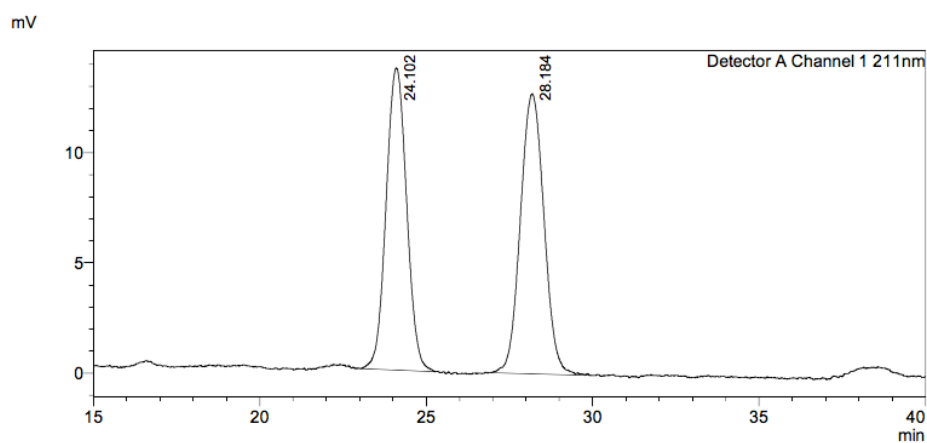
HPLC data for compound **93**: Chiralcel IC (95:5 hexane:IPA, flow rate 1 mLmin<sup>-1</sup>, 211 nm, 30 °C), *t<sub>R</sub>* (A): 24.1 min, *t<sub>R</sub>* (B): 28.1 min; 48:52 ee.

#### <Chromatogram>



Detector A Channel 1 211nm		
Peak#	Ret. Time	Area%
1	24.163	49.593
2	28.191	50.407
Total		100.000

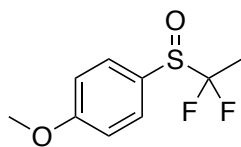
#### <Chromatogram>



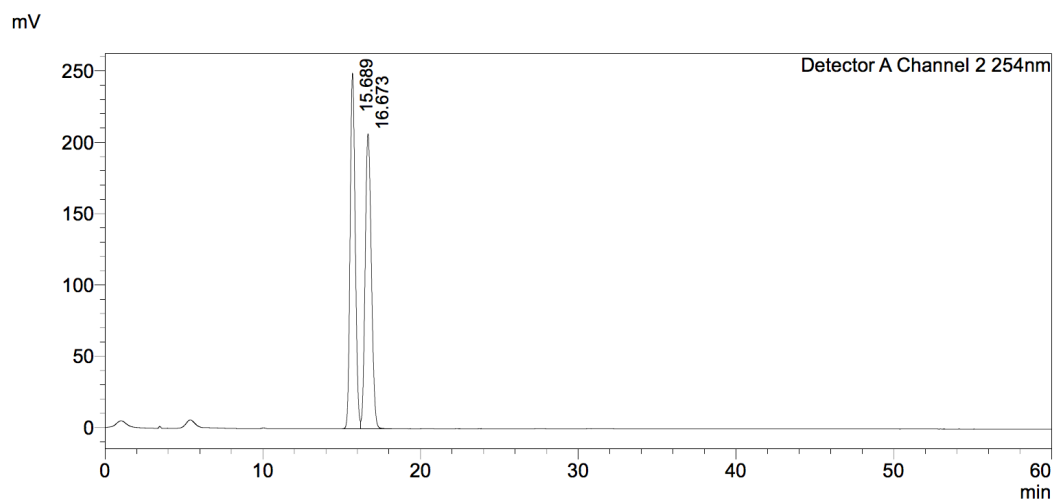
Detector A Channel 1 211nm		
Peak#	Ret. Time	Area%
1	24.102	48.397
2	28.184	51.603
Total		100.000

## 5.4.2 Enantiomeric excess analysis for $\alpha,\alpha$ -difluoroethyl thioethers

### 5.4.2.1 1-((1,1-Difluoroethyl)sulfinyl)-4-methoxybenzene (**112**)

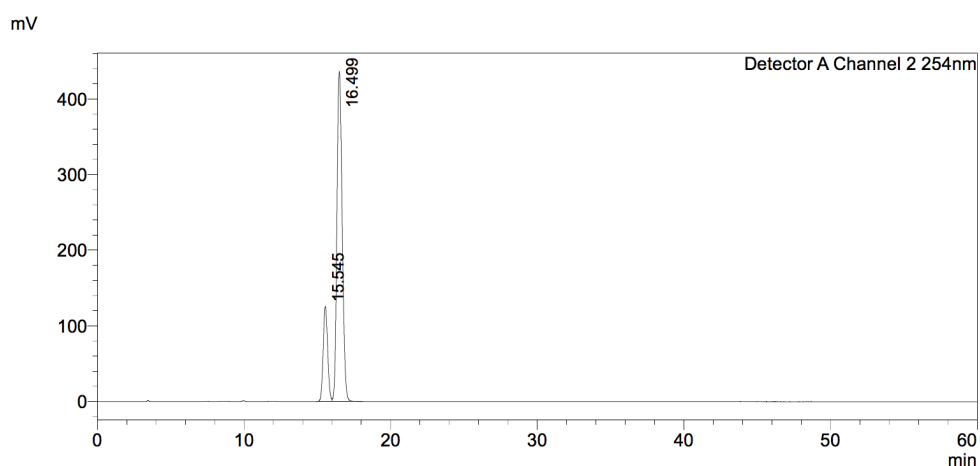


HPLC data for compound **112**: Chiralcel ID (95:5 hexane:IPA, flow rate 1 mLmin<sup>-1</sup>, 254 nm, 30 °C),  $t_R$  (A): 15.6 min,  $t_R$  (B): 16.5 min; 20:80 ee.



#### <Peak Table>

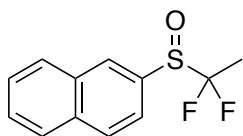
Detector A Channel 2 254nm		
Peak#	Ret. Time	Area%
1	15.689	49.906
2	16.673	50.094
Total		100.000



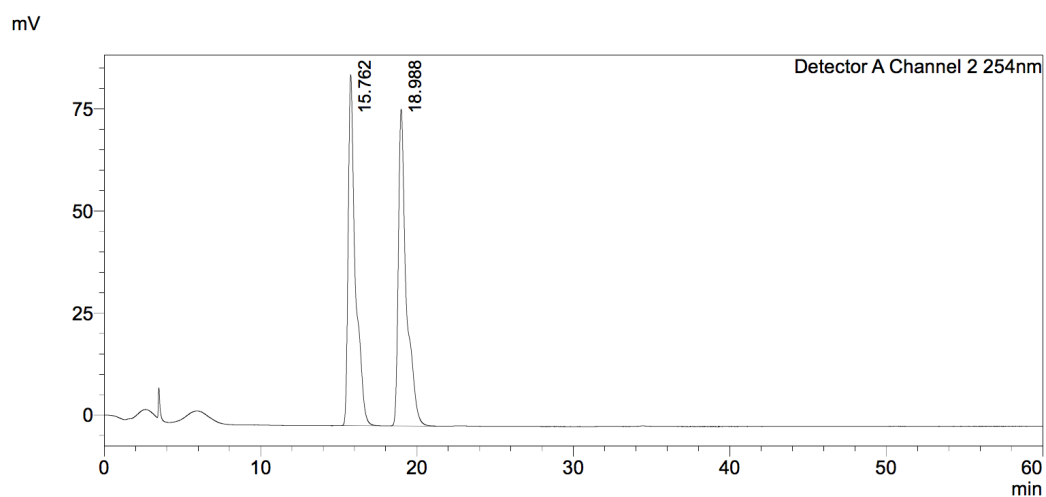
#### <Peak Table>

Detector A Channel 2 254nm		
Peak#	Ret. Time	Area%
1	15.545	19.640
2	16.499	80.360
Total		100.000

#### 5.4.2.2 2-((1,1-Difluoroethyl)sulfinyl)naphthalene (**116**)

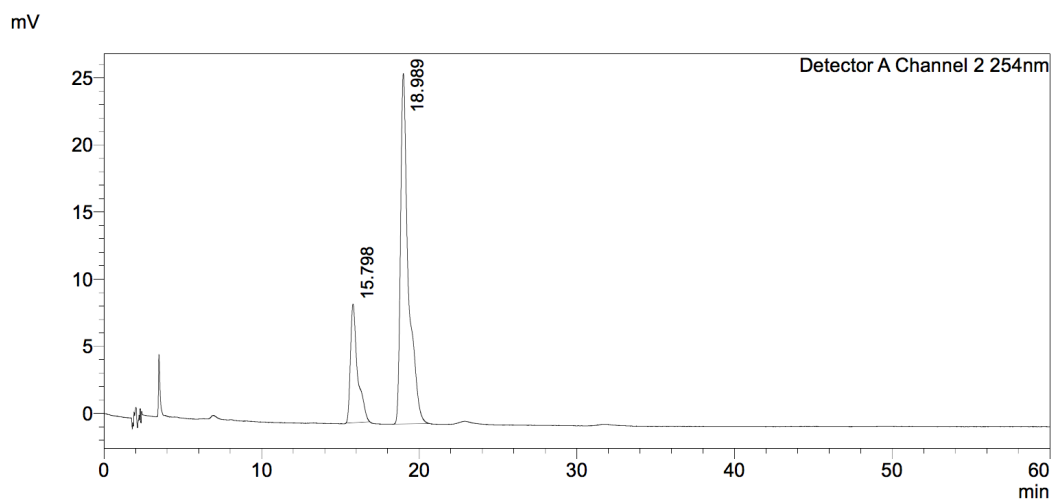


HPLC data for compound **116**: Chiralcel IC (95:5 hexane:IPA, flow rate 1 mLmin<sup>-1</sup>, 254 nm, 30 °C), *t<sub>R</sub>* (A): 18.9 min, *t<sub>R</sub>* (B): 16.5 min; 23:77 ee.



##### <Peak Table>

Detector A Channel 2 254nm		
Peak#	Ret. Time	Area%
1	15.762	49.991
2	18.988	50.009
Total		100.000



##### <Peak Table>

Detector A Channel 2 254nm		
Peak#	Ret. Time	Area%
1	15.798	23.390
2	18.989	76.610
Total		100.000

## 5.5 Estimation of lipophilicities

### 5.5.1 Estimation by $^{19}\text{F}$ NMR

#### 5.5.1.1 General procedures for the estimation of Log P by $^{19}\text{F}$ NMR<sup>10</sup>

The compounds to be measured were added into a pear-bottom flask containing equal amounts of octanol and water. The mixture was stirred at r.t. and the layers allowed to separate overnight.

The NMR samples were prepared carefully to avoid contamination between the phases. A fraction of the upper octanol layer (0.8 mL) was taken with a 1 mL syringe and a needle, and 0.3 mL were discarded to ensure that no aqueous residue was left in the needle. The remaining sample (0.5 mL) was transferred into an NMR tube.  $\text{D}_6$ -acetone (0.1 mL) was added to the tube, and it was stoppered and carefully sealed with Parafilm®. The tube was inverted (20 times) to ensure homogenisation.

For the preparation of the lower aqueous sample, some air was taken with the syringe and needle and it was slowly discharged while passing through the organic layer to avoid accidentally collecting any octanol. The remaining air was slowly discharged into the aqueous phase, and a sample of the aqueous layer was taken (0.8 mL). The needle and syringe were removed as fast as possible, and 0.3 mL of water were discarded to avoid contamination. The needle was also externally wiped with a tissue. The remaining sample (0.5 mL) was added into an NMR tube, followed by  $\text{d}_6$ -acetone (0.1 mL). The tube was stoppered and sealed with Parafilm® prior to inversion (20 times) for homogenisation.

The settings of both the actual NMR experiment and that of the processing of the FID were carried out as follows. The pulse delay time (D1) were evaluated by Linclau such that D1 for octanol should be 30 s and for water, 60 s.<sup>10</sup> The frequency offset point (O1P) should be centred on the fluorine signals, so it will vary for each sample.

All spectra were processed in the same way, using Topspin 3.5pl5. The sequence of commands was as follows: 1. Open spectrum 2. efp 3. Window function → LB (line broadening) = 2 4. efp 5. si = 256 k 6. efp 7. pH 8. bas (Interval baseline correction) 9. .pp → Delete → Integrate → D.

For the spectra analysis, it was first necessary to define the integration ratio ( $\rho$ ) for the aqueous and the octanol phases. In the following equations, 'X' is the compound to be measured and 'ref' is the reference.

$$\rho_{\text{oct}} = \frac{I_{\text{oct}}^{\text{X}}}{I_{\text{oct}}^{\text{ref}}} = \frac{n^{\text{X}}C_{\text{oct}}^{\text{X}}}{n^{\text{ref}}C_{\text{oct}}^{\text{ref}}}$$

$$\rho_{\text{aq}} = \frac{I_{\text{aq}}^{\text{X}}}{I_{\text{aq}}^{\text{ref}}} = \frac{n^{\text{X}}C_{\text{aq}}^{\text{X}}}{n^{\text{ref}}C_{\text{aq}}^{\text{ref}}}$$

Once the integration ratio for both phases has been defined, the partition coefficient ( $P$ ) was calculated as follows:

$$\frac{\rho_{\text{oct}}}{\rho_{\text{aq}}} = \frac{C_{\text{oct}}^{\text{X}}C_{\text{aq}}^{\text{ref}}}{C_{\text{aq}}^{\text{X}}C_{\text{oct}}^{\text{ref}}} = \frac{P^{\text{X}}}{P^{\text{ref}}}$$

$$P^{\text{X}} = P^{\text{ref}} \left( \frac{\rho_{\text{oct}}}{\rho_{\text{aq}}} \right)$$

$$\log P^{\text{X}} = \log P^{\text{ref}} + \log \left( \frac{\rho_{\text{oct}}}{\rho_{\text{aq}}} \right)$$

Since the  $\log P$  of the reference is well known, the only unknown variable is the  $\log P$  of the compound.



### 5.5.1.2 Calibration Experiments

#### 2,2,2-Trifluoroethanol (71) to estimate 1,3-difluoro-2-propanol (72)

2,2,2-Trifluoroethanol (TFE, 5 mg, 3.6  $\mu$ L, **71**) and 1,3-difluoro-2-propanol (5 mg, 4.03  $\mu$ L, **72**) were added into a pear-shaped flask containing octanol (2 mL, HPLC grade) and a stirrer bar. Finally, water was added (2 mL, milliQ). The flask was stirred for 3 hours at room temperature, and the phases were left to separate overnight. NMR samples were prepared as previously described. Full range  $^{19}\text{F}$   $\{^1\text{H}\}$  NMRs were run, the number of scans (NS) were set to the automatic 64 scans, and the O1P was set as -155 Hz. The FID processing was carried out with Topspin as previously described, following the parameters indicated above and integrating each signal manually. This led to the following results:

$$\rho_{\text{water}} = \frac{I^{72}}{I^{71}} = \frac{100.00}{60.24} = 1.62$$

$$\rho_{\text{octanol}} = \frac{I^{72}}{I^{71}} = \frac{27.08}{100.00} = 0.271$$

$$\log P^{71} = +0.36$$

$$\log P^{72} = \log P^{71} + \log \left( \frac{\rho_{\text{oct}}}{\rho_{\text{water}}} \right) = +0.36 + \log \left( \frac{0.271}{1.62} \right) = +0.36 + (-0.79) = -0.43$$

#### TFE (71) to estimate 2-fluorophenol (73)

TFE (5 mg, 3.6  $\mu$ L, **71**) and 2-fluorophenol (5 mg, 3.98  $\mu$ L, **73**) were added into a pear-shaped flask containing octanol (2 mL, HPLC grade) and a stirring bar. MilliQ water was added (2 mL) and the flask was stirred for 3 hours at room temperature. The octanol and aqueous phases were left to separate overnight. A full range  $^{19}\text{F}$   $\{^1\text{H}\}$  NMR was run, with NS = 64 scans and the O1P adjusted to -105 Hz.

$$\rho_{\text{water}} = \frac{I^{73}}{I^{71}} = \frac{1.00}{100.00} = 0.01$$

$$\rho_{\text{octanol}} = \frac{I^{73}}{I^{71}} = \frac{23.03}{100.00} = 0.2303$$

$$\log P^{71} = +0.36$$

$$\log P^{73} = \log P^{71} + \log \left( \frac{\rho_{\text{oct}}}{\rho_{\text{water}}} \right) = +0.36 + \log \left( \frac{0.2303}{0.01} \right) = +0.36 + (1.36) = +1.72$$

### 5.5.1.3 Estimation of Log P values

#### All-*cis* tetrafluorocyclohexylbenzene (**62**)

2-Fluorophenol (5 mg, 3.98  $\mu\text{L}$ , **73**) and all-*cis* tetrafluorocyclohexylbenzene (5.32 mg, **62**) were added into a pear-shaped flask containing octanol (2 mL, HPLC grade) and a stirring bar. MilliQ water was added (2 mL) and the flask was vigorously stirred for 3 hours at room temperature. The octanol and aqueous phases were left to separate overnight. NMR samples were run using the following settings: aqueous sample (NS = 64 scans, D1 = 60 s, O1P = -169 ppm); octanol sample (NS = 64 scans, D1 = 30 s, O1P = -169 ppm), leading to the following calculations and results:

$$\rho_{\text{water}} = \frac{I^{62}}{I^{73}} = \frac{16.65 + 16.71}{100.00} = 0.3336$$

$$\rho_{\text{octanol}} = \frac{I^{62}}{I^{73}} = \frac{95.58 + 90.94}{100.00} = 1.8652$$

$$\log P^{73} = +1.70$$

$$\log P^X = \log P^{\text{ref}} + \log \left( \frac{\rho_{\text{oct}}}{\rho_{\text{water}}} \right) = +1.70 \log \left( \frac{1.8652}{0.3336} \right) = +1.70 + (0.747) = 2.45$$

A second assay was run with 2-fluorophenol (5 mg, 3.98  $\mu\text{L}$ , **73**) and **62** (6.46 mg), obtaining the following results:

$$\rho_{\text{water}} = \frac{I^{62}}{I^{73}} = \frac{19.66 + 16.30}{100.00} = 0.3596$$

$$\rho_{\text{octanol}} = \frac{I^{62}}{I^{73}} = \frac{1.0598 + 0.9873}{1.00} = 2.0471$$

$$\log P^{73} = +1.70$$

$$\log P^{62} = \log P^{73} + \log \left( \frac{\rho_{\text{oct}}}{\rho_{\text{water}}} \right) = +1.70 \log \left( \frac{2.0471}{0.3596} \right) = +1.70 + (0.755) = 2.46$$

### All-*cis* trifluorocyclohexylbenzene (**59**)

2-Fluorophenol (5 mg, 3.98  $\mu$ L, **73**) and all-*cis* trifluorocyclohexylbenzene (5.25 mg, **59**) were added into a pear-shaped flask containing octanol (2 mL, HPLC grade) and a stirrer bar. MilliQ water was added (2 mL) and the flask was vigorously stirred for 3 hours at room temperature. The octanol and aqueous phases were left to separate overnight.

NMR samples were run using the following settings: aqueous sample (NS = 64 scans, D1 = 60 s, O1P = -166 ppm); octanol sample (NS = 64 scans, D1 = 30 s, O1P = -166 ppm), leading to the following calculations:

$$\rho_{\text{water}} = \frac{I^{59}}{I^{73}} = \frac{0.1202 + 0.1039 + 0.1079}{1.00} = 0.332$$

$$\rho_{\text{octanol}} = \frac{I^{59}}{I^{73}} = \frac{0.5156 + 0.5078 + 0.4865}{1.00} = 1.5099$$

$$\log P^{73} = +1.70$$

$$\log P^{59} = \log P^{73} + \log\left(\frac{\rho_{\text{oct}}}{\rho_{\text{water}}}\right) = +1.70 + \log\left(\frac{1.5099}{0.332}\right) = +1.70 + (0.658) = +2.36$$

The experiment was repeated with 2-fluorophenol (5mg, 3.98  $\mu$ L, **73**) and **59** (5 mg), leading to:

$$\rho_{\text{water}} = \frac{I^X}{I^{\text{ref}}} = \frac{0.1266 + 0.0938 + 0.1130}{1.00} = 0.3334$$

$$\rho_{\text{octanol}} = \frac{I^X}{I^{\text{ref}}} = \frac{0.5415 + 0.5365 + 0.5031}{1.00} = 1.5811$$

$$\log P^{\text{ref}} = +1.70$$

$$\log P^X = \log P^{\text{ref}} + \log\left(\frac{\rho_{\text{oct}}}{\rho_{\text{water}}}\right) = +1.70 + \log\left(\frac{1.5811}{0.3334}\right) = +1.70 + (0.675) = +2.38$$

### All-*cis* difluorocyclohexylbenzene (**58**)

2-Fluorophenol (5 mg, 3.98  $\mu$ L, **73**) and all-*cis* difluorocyclohexylbenzene (4.97 mg, **58**) were added into a pear-shaped flask containing octanol (2 mL, HPLC grade) and a stirring bar. MilliQ water was added (2 mL) and the flask was vigorously stirred for 3 hours at room temperature. The octanol and aqueous phases were left to separate overnight. NMR samples were run using the following settings: aqueous sample (NS = 64 scans, D1 = 60 s, O1P = -168 ppm); octanol sample (NS = 64 scans, D1 = 30 s, O1P = -168 ppm). The very low intensity of the fluorine signals in the aqueous phase did not allow the determination of a Log *P* value for this compound.

### All-*cis* tetrafluorocyclohexyl carboxylic acid (**63**)

HPLC grade octanol (2 mL) was added to a pear-shaped flask containing a stirring bar, prior to addition of 2-fluorophenol (2.5 mg, 2  $\mu$ L, **73**) and **63** (1.01 mg). MilliQ grade water was added to the mixture (2 mL), and it was stirred for 3 h at r.t. The phases were left to separate overnight. NMR samples were prepared as previously described, employing the following parameters: aqueous sample (NS = 64 scans, D1 = 60 s, O1P = -170 ppm); octanol sample (NS = 64 scans, D1 = 30 s, O1P = -170 ppm).

$$\rho_{water} = \frac{I^{63}}{I^{73}} = \frac{0.4964 + 0.4384}{1.00} = 0.9348$$

$$\rho_{octanol} = \frac{I^{63}}{I^{73}} = \frac{0.2174 + 0.2053}{1.00} = 0.4227$$

$$\log P^{73} = +1.70$$

$$\log P^{63} = \log P^{73} + \log \left( \frac{\rho_{oct}}{\rho_{water}} \right) = +1.70 + \log \left( \frac{0.4227}{0.9348} \right) = +1.70 + (-0.345) = +1.36$$

### All-*cis* trifluorocyclohexyl carboxylic acid (**60**)

HPLC grade octanol (2 mL) was added to a pear-shaped flask containing a stirring bar, prior to addition of 2-fluorophenol (2.5 mg, 2  $\mu$ L, **73**) and **60** (1.39 mg). MilliQ grade water was added to the mixture (2 mL), and it was stirred for 3 h at r.t. The phases were left to separate overnight. NMR samples were prepared as previously described, employing the following parameters: aqueous sample (NS = 64 scans, D1 = 60 s, O1P = -166 ppm); octanol sample (NS = 64 scans, D1 = 30 s, O1P = -166 ppm).

$$\rho_{water} = \frac{I^{60}}{I^{73}} = \frac{0.2429 + 0.1877 + 0.2679}{1.00} = 0.6985$$

$$\rho_{octanol} = \frac{I^{60}}{I^{73}} = \frac{14.40 + 13.39 + 12.72}{100.00} = 0.4051$$

$$\log P^{73} = +1.70$$

$$\log P^{60} = \log P^{73} + \log \left( \frac{\rho_{oct}}{\rho_{water}} \right) = +1.70 + \log \left( \frac{0.4051}{0.6985} \right) = +1.70 + (-0.237) = +1.46$$

## 5.5.2 Estimation by reverse-phase HPLC

### 5.5.2.1 General procedures for the estimation of Log *P* values by reverse-phase HPLC

The estimation of lipophilicity values was conducted using a Phenomenex Luna C<sub>18</sub> 100A (250 mm × 4.60 mm) 5μ column in a Shimadzu Prominence HPLC. A series of reference compounds were injected (5–10 μL of 0.5 mg/mL in AcCN) and their retention times (*R<sub>t</sub>*) plotted against their Log *P* values to obtain a line equation.

When possible, the value for the references used is the one recommended by Hansch and Leo (1987, 1995) or Sangster (1989).<sup>11,12,13</sup>

The characteristic retention times (*R<sub>t</sub>*) of each of the reference molecules was used to calculate their capacity factor (*k*), using the following equation:

$$\text{Capacity factor (k)} = \frac{\text{Retention Time} - \text{Dead Time of the column}}{\text{Dead Time of the column}}$$

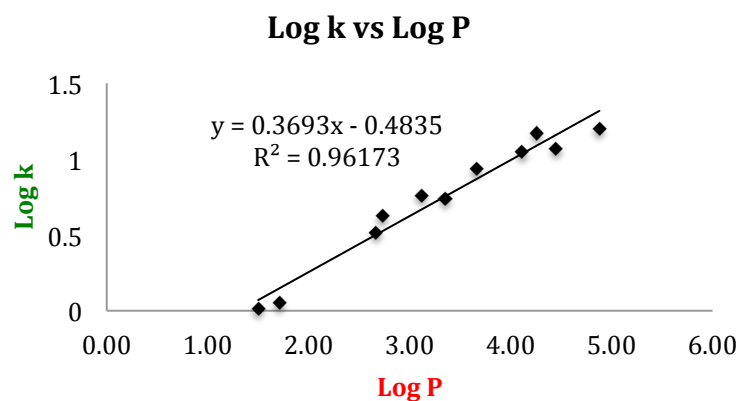
Where the dead time of the column is the time that takes for an unretained molecule (such as the solvent) to pass through the column, and in this case is of 1.97 min.

### 5.5.2.2 Conditions for cyclohexyl derivatives (Section 2.1.4)

The retention times of all the compounds were measured using 60:40 AcCN:water as eluent, both supplemented with 0.05% TFA. The flow rate was of 1 mL/min. The retention time for each reference compound was measured in triplicates, to allow the calculation of an average retention time, which was used to calculate the corresponding capacity factor (k).

Reference	Log P	Rt 1 (min)	Rt 2 (min)	Rt 3 (min)	Average Rt (min)	Capacity factor (k)	Log k
Phenol	1.50	3.99	3.99	3.99	3.99 ± 0.00	1.03	0.013
2-Fluorophenol	1.71	4.18	4.18	4.18	4.18 ± 0.00	1.13	0.051
Benzofuran	2.67	8.42	8.42	8.43	8.42 ± 0.01	3.29	0.52
Toluene	2.73	10.27	10.27	10.27	10.27 ± 0.00	4.23	0.63
<i>o</i> -Xylene	3.12	13.36	13.34	13.37	13.36 ± 0.02	5.80	0.76
Naphthalene	3.35	12.75	12.75	12.78	12.76 ± 0.02	5.50	0.74
Cumene	3.66	18.87	18.88	18.91	18.89 ± 0.02	8.61	0.94
<i>t</i> -Butylbenzene	4.11	23.93	23.86	23.91	23.90 ± 0.04	11.16	1.05
Butylbenzene	4.26	31.10	31.15	31.14	31.13 ± 0.03	14.84	1.17
Anthracene	4.45	25.08	25.07	25.02	25.06 ± 0.04	11.75	1.07
Pyrene	4.88	33.24	33.33	33.17	33.25 ± 0.08	15.92	1.20

The logarithm of each capacity factor was plotted against their literature Log *P* value to afford the equation of a line. For these conditions,  $y = 0.3693x - 0.4835$ , with a correlation coefficient of  $R^2 = 0.9617$ .



The retention time of each of the products to study was measured thrice, and the average retention time used for the calculation of the corresponding capacity factor. The logarithm of the capacity factor was used with the previous equation to yield the Log *P* values of each compound.

Compound	LogP	Rt 1	Rt 2	Rt 3	Average Rt	k	Logk
62	2.58	7.78	7.78	7.72	7.76 ± 0.04	2.95	0.47
59	2.64	8.08	8.08	8.03	8.06 ± 0.03	3.10	0.49
63	1.50	4.28	4.27	4.28	4.28 ± 0.01	1.18	0.07
60	1.51	4.29	4.30	4.30	4.30 ± 0.01	1.19	0.07
58	3.30	12.68	12.68	12.61	12.66 ± 0.05	5.44	0.74
74	2.79	8.87	8.88	8.87	8.87 ± 0.01	3.52	0.55
52	3.76	17.82	17.81	17.71	17.78 ± 0.07	8.05	0.91
75	4.99	46.88	46.53	46.89	46.77 ± 0.24	22.80	1.36
76	2.28	6.45	6.45	6.45	6.45 ± 0.00	2.28	0.36
77	3.23	12.00	11.99	12.00	12.00 ± 0.01	5.11	0.71



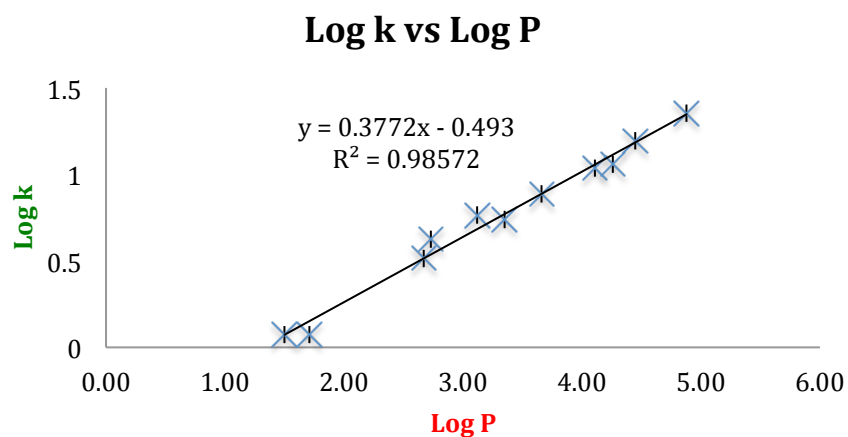
### 5.5.2.2 Conditions for thio- and oxy- ethers (Section 2.2.3) and fluorocyclopropyl derivatives (Section 2.3.3)

The retention times of all the compounds were measured using 60:40 AcCN:water as eluent, with a flow rate of 1 mL/min. No TFA was used for the estimation of these Log *P* values.

The table below displays the experimental results for these conditions, where the retention time for each of the reference compounds was measured three times, in order to calculate an average retention time. The calculated average was used to estimate the capacity factor (*k*) and its logarithm. This log *k* plotted against the known Log *P* value gave an equation of the line for the calculation of the unknown lipophilicities.

Reference	Log <i>P</i>	Rt 1 (min)	Rt 2 (min)	Rt 3 (min)	Average Rt (min)	Capacity factor ( <i>k</i> )	Log <i>k</i>
Phenol	1.50	4.30	4.30	4.26	4.29 ± 0.03	1.18	0.08
2-Fluorophenol	1.71	4.20	4.21	4.50	4.30 ± 0.20	1.18	0.08
Benzofuran	2.67	8.38	8.38	8.38	8.38 ± 0.00	3.26	0.51
Toluene	2.73	10.26	10.27	10.25	10.26 ± 0.01	4.22	0.63
<i>o</i> -Xylene	3.12	13.25	13.28	13.58	13.37 ± 0.30	5.80	0.76
Naphthalene	3.35	12.73	12.73	12.76	12.74 ± 0.02	5.48	0.74
Cumene	3.66	17.06	17.07	17.17	17.10 ± 0.07	7.70	0.89
<i>t</i> -Butylbenzene	4.11	23.45	23.41	23.26	23.37 ± 0.11	10.89	1.04
Butylbenzene	4.26	29.96	30.43	30.45	30.28 ± 0.32	14.41	1.16
Anthracene	4.45	24.46	24.46	24.51	24.48 ± 0.03	11.46	1.06
Pyrene	4.88	32.70	32.62	32.80	32.71 ± 0.09	15.65	1.19

The equation used was  $y = 0.3772x - 0.493$ , with a correlation factor of  $R^2 = 0.9857$ .



The experimental retention times, the calculated capacity factors and their logarithms are shown in the table below, along with the Log  $P$  values (blue), estimated by substitution of the log  $k$  values in the regression line equation.

Compound	LogP	Rt 1	Rt 2	Rt 3	Average Rt	Capacity factor (k)	Logk
PhSCF <sub>3</sub> ( <b>94</b> )	3.70	17.67	17.68	17.68	17.68 ± 0.01	8.00	0.90
PhSCH <sub>2</sub> CH <sub>3</sub> ( <b>95</b> )	3.42	14.32	14.28	14.33	14.31 ± 0.03	6.28	0.80
PhSCF <sub>2</sub> CH <sub>3</sub> ( <b>96</b> )	3.38	13.89	13.88	13.88	13.88 ± 0.01	6.07	0.78
PhSCH <sub>2</sub> CF <sub>3</sub> ( <b>97</b> )	3.30	13.05	13.05	13.03	13.04 ± 0.01	5.64	0.75
PhSCH <sub>3</sub> ( <b>98</b> )	2.87	9.61	9.61	9.62	9.61 ± 0.01	3.89	0.59
PhOCF <sub>3</sub> ( <b>99</b> )	3.35	13.49	13.51	13.49	13.50 ± 0.01	5.87	0.77
PhOCH <sub>2</sub> CH <sub>3</sub> ( <b>100</b> )	2.96	10.42	10.09	10.12	10.21 ± 0.21	4.20	0.62
PhOCF <sub>2</sub> CH <sub>3</sub> ( <b>101</b> )	2.87	9.56	9.57	9.58	9.57 ± 0.01	3.87	0.59
PhOCH <sub>2</sub> CF <sub>3</sub> ( <b>102</b> )	2.83	9.28	9.28	9.48	9.35 ± 0.13	3.76	0.58
PhOCH <sub>3</sub> ( <b>103</b> )	2.51	7.55	7.60	7.57	7.57 ± 0.03	2.85	0.46

Compound	Log <i>P</i>	Rt 1	Rt 2	Rt 3	Average Rt	Capacity factor (k)	Logk
127	4.01	22.45	22.42	22.47	22.45 ± 0.03	10.42	1.02
128	3.42	14.29	14.28	14.31	14.29 ± 0.02	6.27	0.80
129	2.91	9.86	9.86	9.89	9.87 ± 0.02	4.02	0.61
130	2.81	9.21	9.21	9.23	9.22 ± 0.01	3.69	0.57
131	2.79	9.09	9.09	9.11	9.10 ± 0.01	3.63	0.56
126	2.87	9.62	9.63	9.63	9.63 ± 0.01	3.90	0.59
132	3.55	15.76	15.77	15.78	15.77 ± 0.01	7.03	0.85

## 5.6 Enzymatic biotransformations (Chapter 3)

### 5.6.1 Preparation of buffers

#### 5.6.1.1 Phosphate-citrate buffer

The phosphate-citrate buffer was prepared by mixing  $\text{Na}_2\text{HPO}_4$  (0.2 M; 53.7 g of  $\text{Na}_2\text{HPO}_4 \cdot 7\text{H}_2\text{O}$  in 1 L) and citric acid (0.1 M, 19.2 g in 1 L). The volume needed for each pH is described in the following table.<sup>14</sup>

0.2 M $\text{Na}_2\text{HPO}_4$ (mL)	0.1 M citrate (mL)	pH
5.4	44.6	2.6
7.8	42.2	2.8
10.2	39.8	3.0
12.3	37.7	3.2
14.1	35.9	3.4
16.1	33.9	3.6
17.7	32.3	3.8
19.3	30.7	4.0
20.6	29.4	4.2
22.2	27.8	4.4
23.3	26.7	4.6
24.8	25.2	4.8
25.7	24.3	5.0
26.7	23.3	5.2
27.8	22.2	5.4
29.0	21.0	5.6
30.3	19.7	5.8
32.1	17.9	6.0
33.1	16.9	6.2
34.6	15.4	6.4
36.4	13.6	6.6
40.9	9.1	6.8
43.6	6.5	7.0

#### **5.6.1.2 Phosphate buffer (1 M, pH = 7.8)**

Stock solution A (potassium dihydrogenphosphate, 1 M) was mixed with stock solution B (potassium phosphate, dibasic, 1 M) until pH = 7.8. Other phosphate buffers with different pH were prepared from the same stock solutions, in different proportions.

#### **5.6.1.3 Tris·HCl buffer**

A 1M solution of Tris buffer was prepared by addition of Trizma® base (60.6 g) to distilled water (0.3 L), which was stirred for homogenisation. The pH was adjusted by dropwise addition of HCl (conc.), to the required value – as described per each experiment. The volume was adjusted to a total of 0.5 L by addition of distilled water.

The corresponding concentrations of the Tris buffer were obtained by addition of the stock 1 M solution to the required volume of water.

When necessary, the buffer was supplemented with  $\text{Mg}^{2+}$  by addition of the required volume of a 1 M  $\text{MgCl}_2$  stock solution.

#### **5.6.1.4 Buffers for the Purification of FdrC**

1 M Tris·HCl (pH = 8), 1 M imidazole (pH = 8) and 1 M NaCl stock solutions were required for the preparation of the washing buffers employed during FdrC purification with a Ni-NTA column.

**5.6.1.4.1 FdrC ‘Wash Buffer A’ or ‘lysis buffer’:** 20 mM Tris·HCl, 20 mM imidazole and 500 mM NaCl (pH = 8)

Buffer A was prepared by addition of Tris·HCl buffer (10 mL, 1 M, pH = 8), imidazole (10 mL, 1 M, pH = 8) and NaCl (250 mL, 1 M) to a beaker. The solution was homogenised by stirring, and the final volume was adjusted to 500 mL by addition of distilled water.

**5.6.1.4.2 FdrC ‘Wash Buffer B’:** 20 mM Tris·HCl, 50 mM imidazole and 500 mM NaCl (pH = 8)

Buffer B was prepared by addition of Tris·HCl buffer (5 mL, 1 M, pH = 8), imidazole (12.5 mL, 1 M, pH = 8) and NaCl (125 mL, 1 M) to a beaker. The solution was homogenised by stirring, and the final volume was adjusted to 250 mL by addition of distilled water.

**5.6.1.4.3 FdrC ‘Wash Buffer C’ of ‘elution buffer’:** 20 mM Tris·HCl, 400 mM imidazole and 500 mM NaCl (pH = 8)

Buffer C was prepared by addition of Tris·HCl buffer (5 mL, 1 M, pH = 8), imidazole (100 mL, 1 M, pH = 8) and NaCl (125 mL, 1 M) to a beaker. The solution was homogenised by stirring, and the final volume was adjusted to 250 mL by addition of distilled water.

**5.6.1.4.4 FdrC ‘Dialysis buffer’:** 20 mM Tris·HCl, 100 mM NaCl (pH = 8)

Tris·HCl buffer (80 mL, 1 M, pH = 8) and NaCl (400 mL, 1 M) were added to a 5 L beaker and homogenised by stirring. The final volume was adjusted to 4 L by addition of distilled water, and the final pH adjusted to 8.

#### ***5.6.1.5 Buffers for the Purification of Nucleoside Hydrolase***

**5.6.1.5.1 ‘Wash Buffer A’ or ‘lysis buffer’:** 10 mM phosphate buffer, 20 mM imidazole and 150 mM NaCl (pH = 7.8)

Phosphate buffer (5 mL, 1 M, pH = 7.8), imidazole (10 mL, 1 M, pH = 7.8) and NaCl (75 mL, 1 M) were added to a beaker and homogenised. The final volume was adjusted to 500 mL by addition of distilled water.

**5.6.1.5.2 ‘Wash Buffer B’ or ‘elution buffer’:** 10 mM phosphate buffer, 500 mM imidazole and 150 mM NaCl (pH = 7.8)

Phosphate buffer (5 mL, 1 M, pH = 7.8), imidazole (250 mL, 1 M, pH = 7.8) and NaCl (75 mL, 1 M) were added to a beaker and homogenised. The final volume was adjusted to 500 mL by addition of distilled water.

#### **5.6.1.5.3 NH ‘Dialysis buffer’:** 20 mM HEPES, 1 mM CaCl<sub>2</sub>, 150 mM NaCl (pH = 7)

HEPES buffer (80 mL, 1 M, pH = 7), CaCl<sub>2</sub> (4 mL, 1 M) and NaCl (600 mL, 1 M) were added to a 5 L beaker and homogenised. The pH was adjusted to 7.0 and the final volume was adjusted to 4 L by addition of distilled water.

### **5.6.2 Media preparation**

#### **5.6.2.1 Lysogeny Broth Medium (LB or Luria Broth Medium) for *E. coli***

Tryptone (10 g), yeast extract (5 g) and sodium chloride (10 g) were dissolved in distilled water (900 mL). The pH was adjusted to 7.25 by dropwise addition of NaOH (1 M). The appropriate antibiotic (ampicillin for NH and kanamycin for FdrC) was inoculated as specified for each case, and the final volume adjusted to 1000 mL. The medium was sterilised by autoclaving.

#### **5.6.2.1 Lysogeny Broth Agar for *E. coli***

Tryptone (10 g), yeast extract (5 g), sodium chloride (10 g) and agar (15 g) were dissolved in distilled water (900 mL). The pH was adjusted to 7.25 by dropwise addition of NaOH (1 M). The appropriate antibiotic (ampicillin for NH and kanamycin for FdrC) was inoculated as specified for each case, and the final volume adjusted to 1000 mL. The medium was sterilised by autoclaving.

### **5.6.3 Over-expression and purification of nucleoside hydrolase (NH)<sup>6,15</sup>**

An overnight inoculum culture of *E. coli* containing the nucleoside hydrolase synthetic plasmid from *T. vivax*<sup>3</sup> was prepared out of the glycerol stocks in LB with ampicillin (50 mg/mL), and incubated at 37 °C.

Induction cultures were prepared (6 x 1 L) from the overnight inoculum culture, in LB medium supplemented with ampicillin (1 mL of 50 mg/mL per litre of LB). The cultures were incubated at 37 °C until O.D. was approximately 0.6 and then cooled in ice for 30 min. IPTG (0.3 mM) was added, and the cultures were incubated at 25 °C for 24 h. Cells were then harvested by centrifugation.

The cell pellet was resuspended in lysis buffer, containing Roche Complete Protease Inhibitor and deoxyribonuclease (0.1 mg/mL). The sample was sonicated 6 times (15 secs at 10 mA amplitude each time), putting in ice at every interval. Alternatively, best results are achieved using the 'one-shot' cell disrupter. The cells and lysate were kept in ice at all times. The cell lysate was centrifuged (8000 rpm, 4 °C) and filtered through a 0.45 µM filter paper.

The clear lysate was purified by passing through a Ni<sup>2+</sup> column. The lysate was allowed to pass through the column under gravity, leading to the nucleoside hydrolase being bound to the Ni<sup>2+</sup> resin. The column was washed with lysis buffer (2 x 100 mL) and elution buffer (3 x 100 mL).

Samples of the soluble lysate, the insoluble cell debris, the flow-through from the column, and the different washes were taken and prepared to run through an SDS-PAGE gel. NuPage buffer was added to each collected sample. The samples were boiled for 10 min. at 95 °C and centrifuged (13000 rpm, 5 min.). An aliquot (7 µl) of each sample was loaded to an SDS-PAGE gel, along with ladder marker (10 µl), and the gel was run at 200 V for 50 min. The gel was removed from the case and stained with instant blue stain for 40 min.

The first and second washes with buffer C were diluted into an imidazole-free buffer, loaded into dialysis tubing and set to dialyse against the corresponding buffer (20 mM (4-2-hydroxyethyl)-1-piperazineethanesulphonic acid (HEPES), 1 mM CaCl<sub>2</sub>, 150 mM NaCl, pH = 7) at 4 °C overnight.

The protein in the dialysis tubing was transferred to a Vivaspin concentrator (Vivascience) and was concentrated up to approximately 2 mL. The protein concentration was measured using a NanoDrop 1000 spectrophotometer.

#### **5.6.4 Transformation, over-expression and purification of FdrC<sup>2</sup>**

The competent cells (BL21 gold) were thawed on ice for 10 min, and aliquoted (100 µL) into pre-chilled 14-mL BD Falcon polypropylene round-bottom tubes, aliquoting an additional sample for a control of the transformation. The competent cells were gently mixed by tapping of the tubes.



The DNA plasmid (0.5  $\mu$ L) was added to each transformation reaction and swirled gently by tapping. For the control transformation reaction, the pUC18 control plasmid (1  $\mu$ L) was added to the separate 100  $\mu$ L aliquot of the competent cells and swirled gently.

Each reaction was incubated on ice for 30 minutes. Meanwhile, LB medium was pre-heated at 42  $^{\circ}$ C in a heat block. Each transformation reaction was transferred from the ice to the heat block at 42  $^{\circ}$ C for a heat pulse of 20 s, and incubated in ice again for 2 min.

The pre-heated LB medium (0.9 mL) was added into a sterile eppendorf, along with each transformation reaction, and incubated at 37  $^{\circ}$ C for 1 h, shaking at 225 rpm. The cells from each transformation were concentrated by centrifugation, the supernatant being carefully removed with a pipette, and the cells from each transformation reaction were dissolved in minimal amounts of LB medium for plating onto LB agar dishes, which contained kanamycin.

For the pUC18 control transformation, a pool of medium (195  $\mu$ L) was placed on an LB-ampicillin agar plate. An aliquote of the control transformation reaction (5  $\mu$ L) was added to the pool of medium and spread using a sterile spreader.

The plates with the transformation reactions and the control were incubated overnight at 37  $^{\circ}$ C.

To conclude the transformation procedure, a liquid overnight culture was prepared from an isolated *E. coli* colony from the agar plate. For this, the LB medium (15 mL) was added to a 50 mL falcon tube, followed by kanamycin (15  $\mu$ L of a 50 mg/mL solution in water). The isolated *E. coli* colony was scratched off the agar plate with a micropipette tip and swirled into the medium. The culture was incubated at 37  $^{\circ}$ C overnight.

For the over-expression of the cells, the overnight culture (1 mL) was used to inoculate LB medium (1 L) containing kanamycin (1 mL of a 50 mg/ml kanamycin solution) and set to incubate at 37  $^{\circ}$ C until the cell density reached an absorbance at 0.6 at 600 nm. The culture was then cooled on ice for 30 min. A final concentration of 0.3 mM IPTG (from a 1 M stock solution) was added as induction procedure for FdrC over-

expression. The incubation was continued at 25 °C for 24 h. The cells were harvested by centrifugation (3200 rpm for 15 min at 4°C). The supernatant was removed and the cell pellet was frozen at -80 °C.

To prepare glycerol stocks, 1000 µL of the 0.6 OD culture (before the addition of IPTG 1 M) were taken and added to 50% glycerol solution (400 µL). The mixture was stirred with a pipette and aliquoted to eppendorf tubes (200 µL each). The samples were flash-frozen in liquid nitrogen and kept at -80 °C.

Purification was carried out on a Ni<sup>2+</sup> column. For this, the cell pellet was thawed at room temperature, and the cells were resuspended in lysis buffer (100 mL) with 5mg of DNase and 1 protease cocktail pill. The cell suspension was lysed using a one-shot cell disrupter. The lysate was centrifuged (15000 rpm for 20 min) at 4 °C. The soluble fraction was filtered through a 0.45 µm filter and kept on ice. Samples (20 µL) of the soluble and insoluble lysate were reserved for the SDS-PAGE gel.

The Ni<sup>2+</sup> column (4 mL of sepharose-Ni beads) was washed with 3 column-volumes (CV) of water and then equilibrated with buffer A (5 CV). The soluble lysate was allowed to flow-through, reserving a sample (20 µL) for the SDS-PAGE gel. The column was washed with buffer A (2 x 1 CV), and a sample of each reserved on ice. Buffer B (2 x 1 CV) was eluted afterwards, followed by buffer C (3 x 1 CV), collecting samples for the SDS gel.

NuPage buffer was added to each collected sample (see below for the corresponding amounts for the soluble and insoluble aliquots). The samples were boiled for 10 min. at 95 °C and centrifuged (13000 rpm, 5 min.). An aliquot (7 µL) of each sample was loaded to an SDS-PAGE gel, along with ladder marker (10 µL), and the gel was run at 200 V for 50 min. The gel was removed from the case and stained with instant blue stain for 40 min.

The fractions eluted with buffer B and buffer C were diluted with Tris-HCl buffer (50 mL, 20 mM, pH = 8) and loaded into dialysis tubing and set to dialyse against dialysis buffer (20 mM Tris-HCl pH 8.0, 100 mM NaCl) overnight at 4 °C.

The protein was concentrated using a Vivaspin 10 KDa, by centrifuging at 4500 rpm for 30 minutes at 4 °C, and repeating the procedure until a total volume of approximately 2 mL. The enzyme concentration was measured using OD280 nm (Nanodrop).

### 5.6.5 Individual enzymatic assays

#### 5.6.5.1 Incubation of SAM (16) and fluoride with the fluorinase

SAM·2HCl (**16**, 1 mM) and KF (10 mM) were incubated for 24 h at 37 °C with a solution of fluorinase (0.5 mg/mL) in phosphate-citrate buffer (pH = 7.2). The fluorinase was denatured by heating (95 °C, 10 min) and centrifuged (13000 rpm, 5 min). The supernatant was analysed by  $^{19}\text{F}\{^1\text{H}\}$  NMR.  $^{19}\text{F}\{^1\text{H}\}$  NMR (470.40 MHz, deuterium oxide)  $\delta_{\text{F}}$  -231.1 (5'-FDA, **17**).

#### 5.6.5.2 Incubation of 5'-FDA (17) with purine nucleoside phosphorylase (PNP)

5'-FDA (**17**, 1 mM) was added to an eppendorf tube and dissolved in phosphate-citrate buffer (pH = 7.6). PNP (approx. 5 mg, immobilised) was added, and the biotransformation incubated at 37 °C for 16 h. The PNP was centrifuged and the supernatant was analysed by  $^{19}\text{F}\{^1\text{H}\}$  NMR.  $^{19}\text{F}\{^1\text{H}\}$  NMR (470.40 MHz, deuterium oxide)  $\delta$  -230.7 (5-FDRP, **18**).

#### 5.6.5.3 Incubation of 5-FDRP (18) with phytase

5-FDRP (**18**, obtained enzymatically in Section 5.6.5.2) was incubated with phytase (120 mg) in phosphate-citrate buffer (pH = 5.1) at 48 °C for 48 h. The phytase was denatured by heating (100 °C, 10 min) and centrifuged. The supernatant was analysed by  $^{19}\text{F}\{^1\text{H}\}$  NMR.  $^{19}\text{F}\{^1\text{H}\}$  NMR (470.40 MHz, deuterium oxide)  $\delta_{\text{F}}$  -228.3 (5-FDR, **26**), -230.4 (5-FDR, **26**), -230.7 (5-FDRP, **18**), -230.8 (5'-FDA, spiked, **17**).

#### 5.6.5.4 Incubation of 5-FDR (26) with FdrC

Synthetic 5-FDR (**26**, 1 mM) was incubated with FdrC (2.7 mg/mL) in phosphate-citrate buffer (pH = 7.1) at 37 °C for 30 h. The enzyme was denatured by heating (95 °C, 10 min) and centrifuged. The supernatant was analysed by  $^{19}\text{F}\{^1\text{H}\}$  NMR.  $^{19}\text{F}\{^1\text{H}\}$

**NMR** (470.40 MHz, deuterium oxide)  $\delta_F$  -228.8 (5-FDR, **26**), -231.1 (5-FDR, **26**), -233.5 (5'-FHPA, **2**).

#### **5.6.5.5 Incubation of 5'-FDA (**17**) with nucleoside hydrolase (NH)**

5'-FDA (1 mM, **17**) was added to an eppendorf tube and dissolved in phosphate buffer (20 mM, pH = 7.5). NH (200  $\mu$ L, 54.3 mg/mL) was added, and the biotransformation incubated at 37 °C for 24 h. The NH was denatured by heating (95 °C, 10 min) and centrifuged; the supernatant was analysed by  $^{19}\text{F}\{^1\text{H}\}$  NMR.  $^{19}\text{F}\{^1\text{H}\}$  **NMR** (470.40 MHz, deuterium oxide)  $\delta_F$  -228.4 (5-FDR, **26**), -230.7 (5-FDR, **26**).

### **5.6.6 Experiments via phytase**

#### **5.6.6.1 Transformation from SAM (**16**) and fluoride to 5-FHPA (one pH)**

SAM·2HCl (**16**, 140  $\mu$ L, 10 mM) and KF (70  $\mu$ L, 50 mM) were incubated for 24 h at 37 °C with a solution of fluorinase (45  $\mu$ L, 2.2 mg/mL), PNP (40  $\mu$ L, recombinant, expressed in *E. coli*, Sigma-Aldrich), phytase from wheat (40  $\mu$ L, 2.5 mg/mL, Sigma-Aldrich), FdrC (37  $\mu$ L, 2.7 mg/mL, expressed in *E. coli*) and  $\text{NAD}^+$  (70  $\mu$ L, 10 mM) in PBS buffer. The proteins were denatured at 95 °C for 5 min and centrifuged at 13000 rpm for 10 min. A mixture of the supernatant and  $\text{D}_2\text{O}$  was analysed by  $^{19}\text{F}\{^1\text{H}\}$  NMR spectroscopy.  $^{19}\text{F}\{^1\text{H}\}$  **NMR** (470.40 MHz,  $\text{H}_2\text{O}+\text{D}_2\text{O}$ )  $\delta_F$  -228.8 (5-FDR, **26**), -231.1 (5-FDR, **26**), -231.1 (5-FDRP, **18**), -231.2 (5'-FDI, **153**), -231.3 (5'-FDA, **17**), -233.5 (5-FHPA, **2**).

##### **5.6.6.1.1 Control Experiment 1: No fluorinase addition**

SAM·2HCl (**16**, 140  $\mu$ L, 10 mM) and KF (70  $\mu$ L, 50 mM) were incubated for 24 h at 37 °C with a solution of PNP (40  $\mu$ L, recombinant, expressed in *E. coli*, Sigma-Aldrich), phytase from wheat (40  $\mu$ L, 2.5 mg/mL, Sigma-Aldrich), FdrC (37  $\mu$ L, 2.7 mg/mL, expressed in *E. coli*) and  $\text{NAD}^+$  (70  $\mu$ L, 10 mM) in PBS buffer. The proteins were denatured at 95 °C for 5 min and centrifuged at 13000 rpm for 10 min. A mixture of the supernatant and  $\text{D}_2\text{O}$  was analysed by  $^{19}\text{F}\{^1\text{H}\}$  NMR spectroscopy.  $^{19}\text{F}\{^1\text{H}\}$  **NMR** (470.40 MHz,  $\text{H}_2\text{O}+\text{D}_2\text{O}$ )  $\delta_F$  No fluorine peaks.

#### 5.6.6.1.2 Control Experiment 2: No PNP addition

SAM·2HCl (**16**, 140  $\mu$ L, 10 mM) and KF (70  $\mu$ L, 50 mM) were incubated for 24 h at 37 °C with a solution of fluorinase (45  $\mu$ L, 2.2 mg/mL), phytase from wheat (40  $\mu$ L, 2.5 mg/mL, Sigma-Aldrich), FdrC (37  $\mu$ L, 2.7 mg/mL, expressed in *E. coli*) and NAD<sup>+</sup> (70  $\mu$ L, 10 mM) in PBS buffer. The proteins were denatured at 95 °C for 5 min and centrifuged at 13000 rpm for 10 min. A mixture of the supernatant and D<sub>2</sub>O was analysed by <sup>19</sup>F{<sup>1</sup>H} NMR spectroscopy. <sup>19</sup>F{<sup>1</sup>H} NMR (470.40 MHz, H<sub>2</sub>O+D<sub>2</sub>O)  $\delta_F$  -231.3 (5'-FDA, **17**).

#### 5.6.6.1.3 Control Experiment 3: No phytase addition

SAM·2HCl (**16**, 140  $\mu$ L, 10 mM) and KF (70  $\mu$ L, 50 mM) were incubated for 24 h at 37 °C with a solution of fluorinase (45  $\mu$ L, 2.2 mg/mL), PNP (40  $\mu$ L, recombinant, expressed in *E. coli*, Sigma-Aldrich), FdrC (37  $\mu$ L, 2.7 mg/mL, expressed in *E. coli*) and NAD<sup>+</sup> (70  $\mu$ L, 10 mM) in PBS buffer. The proteins were denatured at 95 °C for 5 min and centrifuged at 13000 rpm for 10 min. A mixture of the supernatant and D<sub>2</sub>O was analysed by <sup>19</sup>F{<sup>1</sup>H} NMR spectroscopy. <sup>19</sup>F{<sup>1</sup>H} NMR (470.40 MHz, H<sub>2</sub>O+D<sub>2</sub>O)  $\delta_F$  -228.8 (5-FDR, **26**), -231.1 (5-FDR, **26**), -231.3 (5'-FDA, **17**).

#### 5.6.6.1.4 Control Experiment 4: No FdrC addition

SAM·2HCl (**16**, 140  $\mu$ L, 10 mM) and KF (70  $\mu$ L, 50 mM) were incubated for 24 h at 37 °C with a solution of fluorinase (45  $\mu$ L, 2.2 mg/mL), PNP (40  $\mu$ L, recombinant, expressed in *E. coli*, Sigma-Aldrich), phytase from wheat (40  $\mu$ L, 2.5 mg/mL, Sigma-Aldrich) and NAD<sup>+</sup> (70  $\mu$ L, 10 mM) in PBS buffer. The proteins were denatured at 95 °C for 5 min and centrifuged at 13000 rpm for 10 min. A mixture of the supernatant and D<sub>2</sub>O was analysed by <sup>19</sup>F{<sup>1</sup>H} NMR spectroscopy. <sup>19</sup>F{<sup>1</sup>H} NMR (470.40 MHz, H<sub>2</sub>O+D<sub>2</sub>O)  $\delta_F$  -228.8 (5-FDR, **26**), -231.1 (5-FDR, **26**).

#### 5.6.6.1.5 Control Experiment 5: No NAD<sup>+</sup> addition

SAM·2HCl (**16**, 140  $\mu$ L, 10 mM) and KF (70  $\mu$ L, 50 mM) were incubated for 24 h at 37 °C with a solution of fluorinase (45  $\mu$ L, 2.2 mg/mL), PNP (40  $\mu$ L, recombinant, expressed in *E. coli*, Sigma-Aldrich), phytase from wheat (40  $\mu$ L, 2.5 mg/mL, Sigma-Aldrich) and FdrC (37  $\mu$ L, 2.7 mg/mL, expressed in *E. coli*) in PBS buffer. The

proteins were denatured at 95 °C for 5 min and centrifuged at 13000 rpm for 10 min. A mixture of the supernatant and D<sub>2</sub>O was analysed by <sup>19</sup>F{<sup>1</sup>H} NMR spectroscopy. <sup>19</sup>F{<sup>1</sup>H} NMR (470.40 MHz, H<sub>2</sub>O+D<sub>2</sub>O) δ<sub>F</sub> -228.8 (5-FDR, **26**), -231.1 (5-FDRP, **18**), -231.2 (5'-FDI, **153**), -231.3 (5'-FDA, **17**).

#### **5.6.6.2 Transformation from SAM 16 and fluoride to 5-FHPA 2 (three pHs)**

SAM·2HCl (**16**, 1 mM) and KF (10 mM) were added into an eppendorf tube, followed by phosphate-citrate buffer (pH = 7.2). The mixture was incubated for 24 h at 37 °C and 30 rpm with a solution of fluorinase (0.5 mg/mL) and PNP (approx. 5 mg, immobilised, GSK). The PNP was then filtered, and the fluorinase denatured (95 °C) and centrifuged (13000 rpm, 5 min). The pH of the supernatant was acidified to 5.1 with citric acid (dropwise addition, 1 M), and the phytase from wheat (110 mg) was added. The biotransformation was incubated at 40 °C and 30 rpm for 48 h. The pH was risen to 7.1, causing the phytase to precipitate. The phytase was centrifuged (13000 rpm, 5 min). FdrC (0.5 mg/mL) and NAD<sup>+</sup> (5 mM) were added to the supernatant and incubated at 37 °C for 30 h. FdrC was denatured at 95 °C for 5 min and centrifuged at 13000 rpm for 10 min. A mixture of the supernatant and D<sub>2</sub>O was analysed by <sup>19</sup>F{<sup>1</sup>H} NMR spectroscopy. <sup>19</sup>F{<sup>1</sup>H} NMR (470.40 MHz, H<sub>2</sub>O+D<sub>2</sub>O) δ<sub>F</sub> -230.7 (5-FDRP, **18**), -233.5 (5-FHPA, **2**).

#### **5.6.7 Experiments via NH**

##### **5.6.7.1 Transformation from SAM (16) and fluoride to 5-FHPA (2) in PBS buffer**

SAM·2HCl (**16**, 1 mM) and KF (10 mM) were incubated for 48 h at 37 °C with a solution of fluorinase (1.7 mg/mL), NH (3.6 mg/mL), FdrC (2.7 mg/mL) and NAD<sup>+</sup> (5 mM) in PBS buffer (pH = 7.4). The proteins were denatured at 95 °C for 5 min and centrifuged at 13000 rpm for 10 min. A mixture of the supernatant and D<sub>2</sub>O was analysed by <sup>19</sup>F{<sup>1</sup>H} NMR spectroscopy. <sup>19</sup>F{<sup>1</sup>H} NMR (470.40 MHz, H<sub>2</sub>O+D<sub>2</sub>O) δ<sub>F</sub> -228.4 (5-FDR, **26**), -230.7 (5-FDR, **26**), -230.8 (5'-FDI, **153**), -231.1 (5'-FDA, **17**), -233.2 (5-FHPA, **2**).

#### ***5.6.7.2 Transformation from SAM (16) and fluoride to 5-FHPA (2) in phosphate buffer***

SAM·2HCl (**16**, 1 mM) and KF (10 mM) were incubated for 24 h at 37 °C with a solution of fluorinase (0.5 mg/mL), NH (1.5 mg/mL), FdrC (37 µL, 2.7 mg/mL, expressed in *E. coli*) and NAD<sup>+</sup> (70 µL, 10 mM) in phosphate buffer (pH = 8). The biotransformation was supplemented with MgCl<sub>2</sub> (2.7 mg/mL). The enzymes were denatured at 95 °C for 5 min and centrifuged at 13000 rpm for 10 min. A mixture of the supernatant and D<sub>2</sub>O was analysed by <sup>19</sup>F{<sup>1</sup>H} NMR spectroscopy. **<sup>19</sup>F{<sup>1</sup>H} NMR** (470.40 MHz, H<sub>2</sub>O+D<sub>2</sub>O) δ<sub>F</sub> -233.2 (5-FHPA, **2**).

#### ***5.6.7.3 Transformation from SAM (16) and fluoride to 5-FHPA (2) in Tris·HCl buffer***

SAM·2HCl (**16**, 1 mM) and KF (10 mM) were incubated for 24 h at 37 °C with a solution of fluorinase (0.5 mg/mL), NH (2.7 mg/mL), FdrC (2.7 mg/mL) and NAD<sup>+</sup> (10 mM) in Tris·HCl buffer (25 mM, pH = 7.8). The biotransformation was supplemented with MgCl<sub>2</sub> (1 mM). The enzymes were denatured at 95 °C for 5 min and centrifuged at 13000 rpm for 10 min. A mixture of the supernatant and D<sub>2</sub>O was analysed by <sup>19</sup>F{<sup>1</sup>H} NMR spectroscopy. **<sup>19</sup>F{<sup>1</sup>H} NMR** (470.40 MHz, H<sub>2</sub>O+D<sub>2</sub>O) δ<sub>F</sub> -231.1 (5'-FDA, **17**), -233.2 (5-FHPA, **2**).

### **5.6.8 Biotransformations with LDH as recycling agent for NAD<sup>+</sup>**

#### ***5.6.8.1 Biotransformation of fluoropyruvate (162) with LDH***

Fluoropyruvate (**162**, 1 mM) was incubated with LDH from rabbit muscle (5 µL, Sigma Aldrich) and NADH (1 mM) in phosphate buffer (20 mM, pH = 7.5) for 18 h at 37 °C. The enzyme was denatured by heating (95 °C, 10 min) and centrifuged. Analysis of the supernatant by <sup>19</sup>F{<sup>1</sup>H} NMR indicated full conversion into (2*R*)-3-fluorolactate (**163**). **<sup>19</sup>F{<sup>1</sup>H} NMR** (470.40 MHz, H<sub>2</sub>O+D<sub>2</sub>O) δ<sub>F</sub> -228.7 (**163**).

#### ***5.6.8.2 Biotransformation of SAM (16) and KF to 5-FHPA (2) via LDH recycling***

SAM·2HCl (**16**, 1 mM) and KF (10 mM) were incubated for 24 h at 37 °C with a solution of fluorinase (0.5 mg/mL), NH (2.7 mg/mL), FdrC (2.7 mg/mL), pyruvate

(1.5 mM), LDH (5  $\mu$ L), and  $\text{NAD}^+$  (0.5 mM) in Tris·HCl buffer (25 mM, pH = 7.8). The biotransformation was supplemented with  $\text{MgCl}_2$  (1 mM). The enzymes were denatured at 95  $^{\circ}\text{C}$  for 5 min and centrifuged at 13000 rpm for 10 min. A mixture of the supernatant and  $\text{D}_2\text{O}$  was analysed by  $^{19}\text{F}\{^1\text{H}\}$  NMR spectroscopy.  $^{19}\text{F}\{^1\text{H}\}$  NMR (470.40 MHz,  $\text{H}_2\text{O}+\text{D}_2\text{O}$ )  $\delta_{\text{F}}$  -231.1 (5'-FDA, **17**), -233.2 (5-FHPA, **2**).

#### ***5.6.8.3 Biotransformation of 5-FDR (26) via LDH recycling***

5-FDR (**26**, 1 mM) was incubated for 20 h at 48  $^{\circ}\text{C}$  with a solution of FdrC (2.7 mg/mL), pyruvate (1.5 mM), LDH (5  $\mu$ L), and  $\text{NAD}^+$  (0.5 mM) in Tris·HCl buffer (25 mM, pH = 7.5). The biotransformation was supplemented with  $\text{MgCl}_2$  (1 mM). The enzymes were denatured at 95  $^{\circ}\text{C}$  for 5 min and centrifuged at 13000 rpm for 10 min. A mixture of the supernatant and  $\text{D}_2\text{O}$  was analysed by  $^{19}\text{F}\{^1\text{H}\}$  NMR spectroscopy.  $^{19}\text{F}\{^1\text{H}\}$  NMR (470.40 MHz,  $\text{H}_2\text{O}+\text{D}_2\text{O}$ )  $\delta_{\text{F}}$  -228.5 (5-FDR, **26**), -230.7 (5-FDR, **26**), -233.2 (5-FHPA, **2**).

#### ***5.6.8.4 Preparative scale biotransformation of 5'-FDA (17) via LDH recycling***

5'-FDA (**17**, 1 mM) was incubated for 20 h at 37  $^{\circ}\text{C}$  with a solution of NH (3.0 mg/mL), FdrC (2.7 mg/mL), pyruvate (1.5 mM), LDH (5  $\mu$ L), and  $\text{NAD}^+$  (0.5 mM) in Tris·HCl buffer (25 mM, pH = 7.5). The biotransformation was supplemented with  $\text{MgCl}_2$  (1 mM). The enzymes were denatured at 95  $^{\circ}\text{C}$  for 5 min and centrifuged at 13000 rpm for 10 min. A mixture of the supernatant and  $\text{D}_2\text{O}$  was analysed by  $^{19}\text{F}\{^1\text{H}\}$  NMR spectroscopy.  $^{19}\text{F}\{^1\text{H}\}$  NMR (470.40 MHz,  $\text{H}_2\text{O}+\text{D}_2\text{O}$ )  $\delta_{\text{F}}$  -233.2 (5-FHPA, **2**).

### **5.6.9 Biotransformations with FMN as recycling agent for $\text{NAD}^+$**

#### ***5.6.9.1 Biotransformation of SAM (16) and KF via FMN recycling (analytical scale)***

SAM·2HCl (**16**, 1 mM) and KF (10 mM) were incubated for 18 h at 37  $^{\circ}\text{C}$  with a solution of fluorinase (0.5 mg/mL), NH (2.7 mg/mL), FdrC (2.7 mg/mL), FMN (12 mM), and  $\text{NAD}^+$  (0.1 mM) in Tris·HCl buffer (25 mM, pH = 7.8). The biotransformation was supplemented with  $\text{MgCl}_2$  (1 mM). The enzymes were denatured at 95  $^{\circ}\text{C}$  for 5 min and centrifuged at 13000 rpm for 10 min. A mixture of the supernatant and  $\text{D}_2\text{O}$  was analysed by  $^{19}\text{F}\{^1\text{H}\}$  NMR spectroscopy.  $^{19}\text{F}\{^1\text{H}\}$  NMR



(470.40 MHz, H<sub>2</sub>O+D<sub>2</sub>O)  $\delta_F$  -231.1 (5'-FDA, **17**), -233.2 (5-FHPA, **2**), minimal traces of both.

#### ***5.6.9.2 Biotransformation of 5'-FDA (17) via FMN recycling (analytical scale)***

5'-FDA (**17**, 1 mM) was incubated for 20 h at 37 °C with a solution of NH (3.0 mg/mL), FdrC (2.7 mg/mL), FMN (12 mM), and NAD<sup>+</sup> (0.1 mM) in Tris·HCl buffer (25 mM, pH = 7.5). The biotransformation was supplemented with MgCl<sub>2</sub> (1 mM). The enzymes were denatured at 95 °C for 5 min and centrifuged at 13000 rpm for 10 min. A mixture of the supernatant and D<sub>2</sub>O was analysed by <sup>19</sup>F{<sup>1</sup>H} NMR spectroscopy. <sup>19</sup>F{<sup>1</sup>H} NMR (470.40 MHz, H<sub>2</sub>O+D<sub>2</sub>O)  $\delta_F$  -233.2 (5-FHPA, **2**).

#### ***5.6.9.3 Biotransformation of SAM (16) and KF via FMN recycling (preparative scale)***

SAM·2HCl (**16**, 5 mg, 1 mM) and KF (10 mM) were incubated for 48 h at 37 °C with a solution of fluorinase (2 mg/mL), NH (3.0 mg/mL), FdrC (3.0 mg/mL), FMN (12 mM), and NAD<sup>+</sup> (0.1 mM) in Tris·HCl buffer (25 mM, pH = 7.8). The biotransformation was supplemented with MgCl<sub>2</sub> (0.5 mM), and was adjusted to a total volume of 5 mL. The enzymes were denatured at 95 °C for 5 min and centrifuged at 13000 rpm for 10 min. A mixture of the supernatant and D<sub>2</sub>O was analysed by <sup>19</sup>F{<sup>1</sup>H} NMR spectroscopy. <sup>19</sup>F{<sup>1</sup>H} NMR (470.40 MHz, H<sub>2</sub>O+D<sub>2</sub>O)  $\delta_F$  -231.1 (5'-FDA, **17**).

#### ***5.6.9.4 Biotransformation of 5'-FDA (17) via FMN recycling (preparative scale)***

5'-FDA (**17**, 10 mg) was incubated for 20 h at 37 °C with a solution of NH (3.0 mg/mL), FdrC (2.7 mg/mL), FMN (12 mM), and NAD<sup>+</sup> (0.5 mM) in Tris·HCl buffer (25 mM, pH = 7.5). The biotransformation was supplemented with MgCl<sub>2</sub> (1 mM), and adjusted to a total volume of 5 mL. The enzymes were denatured at 95 °C for 5 min and centrifuged at 13000 rpm for 10 min. A mixture of the supernatant and D<sub>2</sub>O was analysed by <sup>19</sup>F{<sup>1</sup>H} NMR spectroscopy. <sup>19</sup>F{<sup>1</sup>H} NMR (470.40 MHz, H<sub>2</sub>O+D<sub>2</sub>O)  $\delta_F$  -233.2 (5-FHPA, **2**).

#### **4.6.9.5 Biotransformation of 5-FDR (26) via FMN recycling (preparative scale)**

5-FDR (**26**, 10 mg, 1 mM) was incubated for 48 h at 48 °C with a solution of FdrC (3.0 mg/mL), FMN (12 mM) and NAD<sup>+</sup> (0.1 mM) in Tris·HCl buffer (25 mM, pH = 7.5). The biotransformation was supplemented with MgCl<sub>2</sub> (0.5 mM), and was adjusted to a total volume of 5 mL. The enzymes were denatured at 95 °C for 5 min and centrifuged at 13000 rpm for 10 min. A mixture of the supernatant and D<sub>2</sub>O was analysed by <sup>19</sup>F{<sup>1</sup>H} NMR spectroscopy. **<sup>19</sup>F{<sup>1</sup>H} NMR** (470.40 MHz, H<sub>2</sub>O+D<sub>2</sub>O) δ<sub>F</sub> -228.5 (5-FDR, **26**), -230.7 (5-FDR, **26**), -233.2 (5-FHPA, **2**).

#### **5.6.10 Scale-up of the biosynthesis of 5-FHPA (2) from 5'-FDA (17)**

5'-FDA (**17**, 50 mg, 27 mM) was incubated for 1 week at 37 °C with a solution of NH (4.0 mg/mL), FdrC (4.0 mg/mL), FMN (40 mM), and NAD<sup>+</sup> (0.5 mM) in Tris·HCl buffer (25 mM, pH = 7.5). The biotransformation was supplemented with MgCl<sub>2</sub> (0.5 mM), and the total volume adjusted to 7 mL. The enzymes were denatured at 95 °C for 5 min and centrifuged at 13000 rpm for 10 min. A mixture of the supernatant and D<sub>2</sub>O was analysed by <sup>19</sup>F{<sup>1</sup>H} NMR spectroscopy. **<sup>19</sup>F{<sup>1</sup>H} NMR** (470.40 MHz, H<sub>2</sub>O+D<sub>2</sub>O) δ<sub>F</sub> -228.5 (5-FDR, **26**), -230.7 (5-FDR, **26**), -233.2 (5-FHPA, **2**).

## **5.7 Fluorometabolite studies in *Streptomyces* sp. MA37 (Chapter 4)**

### **5.7.1 Media preparation**

#### ***5.7.1.1 Yeast-Malt Broth (ISP Medium N° 2) for *S. sp. MA37****

Peptic digest of animal tissue (5 g), yeast extract (3 g), malt extract (3 g) and dextrose (10 g) were added to a beaker containing a stirring bar, and dissolved in tap water (800 mL).<sup>\*</sup> The pH was adjusted to 7.0 by dropwise addition of HCl (1 M) while stirring vigorously. The solution was transferred to a measuring cylinder, the total volume was adjusted to 1 L, and the mixture transferred into a Duran bottle for sterilisation in the same day. The sterile medium was stored at 4 °C.

<sup>\*</sup> This medium was supplemented with KF (0.5 M) to a final concentration of 2.5 mM when necessary, as indicated in each experiment. Supplementation occurred prior to the pH adjustment.

#### ***5.7.1.2 Yeast-Malt Agar (YM Agar, ISP Medium N° 2) for *S. sp. MA37****

Peptic digest of animal tissue (5 g), yeast extract (3 g), malt extract (3 g), dextrose (10 g) and agar (20 g) were added to a beaker containing a stirring bar, and dissolved in tap water (800 mL).<sup>\*</sup> The pH was adjusted to 7.0 by dropwise addition of HCl (1 M) while stirring vigorously. The solution was transferred to a measuring cylinder, the total volume was adjusted to 1 L, and the mixture transferred into a Duran bottle for sterilisation in the same day. The sterile medium was stored at 4 °C.

<sup>\*</sup> This medium was supplemented with KF (0.5 M) to a final concentration of 2.5 mM when necessary, as indicated in each experiment. Supplementation occurred prior to the pH adjustment.

#### ***5.7.1.3 Fermentation culture for *S. cattleya****

Ion solution (150 mL), filtered carbon solution (75 mL), sterile phosphate buffer (75 mL, 150 mM, pH = 7) and sterile KF (3 mL, 0.5 M) were mixed with sterile ultra-pure water (450 mL).

The **ion solution** (trace elements solution) was prepared with ammonium chloride (6.75 g), sodium chloride (2.25 g), magnesium sulfate heptahydrate (2.25 g), calcium carbonate (1.13 g), iron sulfate heptahydrate (113 mg), cobalt chloride hexahydrate (45 mg) and zinc sulfate heptahydrate (45 mg) in ultra-pure water (900 mL).

The filtered **carbon source solution** was prepared by dissolving glycerol (45 g), monosodium glutamate (22.5 g), *myo*-inositol (1.8 g) and *p*-aminobenzoic acid (450 µL, 1 mg/mL solution) in ultra-pure water (900 mL). The solution was sterilised by filtration into a pre-sterilised Schott bottle.

#### **5.7.1.4 Growth of *S. cattleya* on solid media**

Soybean flour (2% w/v), mannitol (2% w/v), and agar (1.5% w/v) were dissolved in tap water and sterilised before use.

#### **5.7.1.5 TSBY Rich Medium**

Tryptic soy broth (30 g), sucrose (103 g) and yeast extract (5 g) were dissolved in distilled water (800 mL) in a beaker, and the mixture was stirred until complete homogenisation. The solution was transferred to a measuring cylinder, and the total volume was adjusted to 1 L. The medium was autoclaved in a glass bottle, and stored at 4 °C.

#### **5.7.2 Preparation of cell-free extracts (CFE) from *S. sp.* MA37**

A liquid culture of *S. sp.* MA37 was prepared using ISP2 media, following the usual procedure. After 8 days of fermentation, the cells were centrifuged (r.t., 13000 rpm, 20 minutes). The cell pellets were washed with phosphate buffer (50 mM, pH 7.0) and centrifuged. The washing procedure was repeated three times. The resultant pellet was stored at -80 °C. A sample of the frozen pellet was thawed and re-suspended in phosphate buffer (200 mM, pH 6.8), keeping a concentration of 0.2 g of wet pellet per mL of buffer solution. The cells were disrupted by sonication (60% duty cycle for 30 – 60 seconds), repeating the procedure three times. Cell debris was removed by centrifugation (r.t., 13000 rpm, 30 minutes), and the supernatant was used as a CFE for incubation experiments. CFEs were supplemented with or without the different target

molecules at 37 °C between 6 and 22 hours. All the CFE experiments were prepared to a final concentration of 3 mM of substrate, unless stated otherwise. After the incubation, the protein was precipitated by heating at 95 °C for 5-10 min, and then removed by centrifugation (13000 rpm, 20 minutes). The results were analysed by  $^{19}\text{F}$ -NMR.

#### **5.7.2.1 CFE incubation of SAM (16) and KF**

SAM (**16**, 300  $\mu\text{L}$ , 10 mM) and KF (60  $\mu\text{L}$ , 50 mM) were incubated for 16 h with CFE (640  $\mu\text{L}$ , 0.1 g/mL) of *S. sp.* MA37. The sample was then heated (95 °C, 5 min), and the denatured protein removed by centrifugation (13000 rpm, 20 minutes). A mixture of the supernatant (650  $\mu\text{L}$ ) and  $\text{D}_2\text{O}$  (100  $\mu\text{L}$ ) was analysed by  $^{19}\text{F}\{^1\text{H}\}$  NMR spectroscopy.  $^{19}\text{F}\{^1\text{H}\}$  NMR (470.40 MHz,  $\text{H}_2\text{O}+\text{D}_2\text{O}$ )  $\delta_{\text{F}}$  -119.4 (fluoride ion).

#### **5.7.2.2 CFE incubation of 5'-FDA (17)**

5'-FDA (**17**, 300  $\mu\text{L}$ , 10 mM) was incubated for 16 h with a CFE (700  $\mu\text{L}$ , 0.1 g/mL) of *S. sp.* MA37. The sample was then heated (95 °C, 5 min), and the denatured protein removed by centrifugation (13000 rpm, 20 minutes). A mixture of the supernatant (650  $\mu\text{L}$ ) and  $\text{D}_2\text{O}$  (100  $\mu\text{L}$ ) was analysed by  $^{19}\text{F}\{^1\text{H}\}$  NMR spectroscopy.  $^{19}\text{F}\{^1\text{H}\}$  NMR (470.40 MHz,  $\text{H}_2\text{O}+\text{D}_2\text{O}$ )  $\delta_{\text{F}}$  -230.7 (5'-FDRP, **18**), -230.9 (5'-FDI, **153**), -231.0 (5'-FDA, **17**) and -231.4 (unknown).

#### **5.7.2.3 CFE incubation of 5'-FDI (153)**

5'-FDI (**153**) was incubated for 20 h with a CFE (700  $\mu\text{L}$ , 0.2 g/mL) of *S. sp.* MA37. The sample was then heated (95 °C, 5 min), and the denatured protein removed by centrifugation (13000 rpm, 20 minutes). A mixture of the supernatant (650  $\mu\text{L}$ ) and  $\text{D}_2\text{O}$  (100  $\mu\text{L}$ ) was analysed by  $^{19}\text{F}\{^1\text{H}\}$  NMR spectroscopy.  $^{19}\text{F}\{^1\text{H}\}$  NMR (470.40 MHz,  $\text{H}_2\text{O}+\text{D}_2\text{O}$ )  $\delta_{\text{F}}$  -230.9 (5'-FDI, **153**).

#### **5.7.2.4 CFE incubation of 5-FDR (26)**

5-FDR (**26**, 300  $\mu\text{L}$ , 10 mM) was incubated for 16 h with a CFE (700  $\mu\text{L}$ , 0.1 g/mL) of *S. sp.* MA37. The sample was then heated (95 °C, 5 min), and the denatured protein

removed by centrifugation (13000 rpm, 20 minutes). A mixture of the supernatant (650  $\mu$ L) and D<sub>2</sub>O (100  $\mu$ L) was analysed by  $^{19}\text{F}\{^1\text{H}\}$  NMR spectroscopy.  $^{19}\text{F}\{^1\text{H}\}$  NMR (470.40 MHz, H<sub>2</sub>O+D<sub>2</sub>O)  $\delta_{\text{F}}$  -228.2 (unknown), -228.3 (unknown), -228.8 (5-FDR, **26**), -228.8 (unknown), -230.8 (unknown), -231.0 (5-FDR, **26**), -231.4 (unknown), -233.1 (unknown) and -233.5 (5-FHPA, **2**).

### 5.7.3 Preparation of a cell-free extract of *S. cattleya*

A liquid culture of *S. cattleya* was prepared inoculating a sample of the organism stocked in glycerol stock into a culture of *S. cattleya* fermentation media. After 8 days of fermentation, the cells were centrifuged (r.t., 13000 rpm, 20 min). The cell pellets were washed with phosphate buffer (50 mM, pH 7.0). The suspension was centrifuged, and the pellet was stored at -80 °C. The pellet was thawed and re-suspended in phosphate buffer (200 mM, pH 6.8), keeping a concentration of 0.2 g of wet pellet per mL of buffer solution. The cells were disrupted by sonication (60% duty cycle for 30 – 60 seconds), repeating the procedure three times. Cell debris was removed by centrifugation (r.t., 13000 rpm, 30 min). The supernatant was used as a CFE for incubation experiments. CFE were supplemented with or without the different target molecules at 37 °C for 16 h. All the CFEs experiments were prepared to a final concentration of 3 mM of substrate, unless stated otherwise. After the incubation, protein was precipitated by heating at 95 °C for 5-10 min, and removed by centrifugation (13000 rpm, 20 minutes). Results were analysed by  $^{19}\text{F}\{^1\text{H}\}$  NMR spectroscopy.

#### 5.7.3.1 CFE incubation of SAM (16) and KF

SAM (**16**, 300  $\mu$ L, 10 mM) and KF (60  $\mu$ L, 50 mM) were incubated 16 h with a CFE (640  $\mu$ L, 0.1 g/mL) of *S. cattleya*. The sample was then heated (95 °C, 5 min), and the denatured protein removed by centrifugation (13000 rpm, 20 min). A mixture of the supernatant (650  $\mu$ L) and D<sub>2</sub>O (100  $\mu$ L) was analysed by  $^{19}\text{F}\{^1\text{H}\}$  NMR spectroscopy.  $^{19}\text{F}\{^1\text{H}\}$  NMR (470.40 MHz, H<sub>2</sub>O+D<sub>2</sub>O)  $\delta_{\text{F}}$  -231.2 (4-FT, **4**).

### 5.7.3.2 CFE incubation of 5'-FDA (17)

5'-FDA (**17**, 300  $\mu$ L, 10 mM) was incubated for 16 h with a CFE (700  $\mu$ L, 0.1 g/mL) of *S. cattleya*. The sample was then heated (95  $^{\circ}$ C, 5 min), and the denatured protein removed by centrifugation (13000 rpm, 20 min). A mixture of the supernatant (650  $\mu$ L) and D<sub>2</sub>O (100  $\mu$ L) was analysed by  $^{19}\text{F}\{^1\text{H}\}$  NMR spectroscopy.  $^{19}\text{F}\{^1\text{H}\}$  NMR (470.40 MHz, H<sub>2</sub>O+D<sub>2</sub>O)  $\delta_{\text{F}}$  -217.2 (FAc, **1**), -231.0 (unknown), -231.2 (4-FT, **4**).

### 5.7.4 5'-FDA (17) enzymatic transformations (Section 4.4.1)

#### 5.7.4.1 5'-FDA (17) reaction with NH

A mixture of 5'-FDA (**17**, 500  $\mu$ L, 10 mM), phosphate buffer (20  $\mu$ L, 1M, pH 8.0), NH (100  $\mu$ L, 10 mg/mL) and sterile water (480  $\mu$ L) was incubated for 16 h. The protein was denatured at 95  $^{\circ}$ C for 5 min and centrifuged at 13000 rpm for 10 min. A mixture of the supernatant and D<sub>2</sub>O was analysed by  $^{19}\text{F}\{^1\text{H}\}$  NMR spectroscopy.  $^{19}\text{F}\{^1\text{H}\}$  NMR (470.40 MHz, H<sub>2</sub>O+D<sub>2</sub>O)  $\delta_{\text{F}}$  -228.8 (5-FDR, **26**), -231.1 (5-FDR, **26**), -231.3 (5'-FDI, **153**), -231.5 (5'-FDA, **17**).

#### 5.7.4.2 5'-FDA (17) reaction with ADA

A mixture of 5'-FDA (**17**, 500  $\mu$ L, 10 mM), phosphate buffer (20  $\mu$ L, 1M, pH 7.5), ADA (365  $\mu$ L) and sterile water (144  $\mu$ L) was incubated for 16 h. The protein was denatured at 95  $^{\circ}$ C for 5 min and centrifuged at 13000 rpm for 10 min. A mixture of the supernatant and D<sub>2</sub>O was analysed by  $^{19}\text{F}\{^1\text{H}\}$  NMR spectroscopy.  $^{19}\text{F}\{^1\text{H}\}$  NMR (470.40 MHz, H<sub>2</sub>O+D<sub>2</sub>O)  $\delta_{\text{F}}$  -228.8 (5-FDR, **26**), -231.1 (5-FDR, **26**), -231.2 (5'-FDI, **153**), -231.3 (5'-FDA, **17**), -231.5 (unknown), -231.7 (unknown).

#### 5.7.4.3 5'-FDA (17) reaction with PNP

A mixture of 5'-FDA (**17**, 500  $\mu$ L, 10 mM), phosphate buffer (50  $\mu$ L, 1M, pH 7.8), PNP (200  $\mu$ L) and sterile water (250  $\mu$ L) was incubated for 16 h. The protein was denatured at 95  $^{\circ}$ C for 5 min and centrifuged at 13000 rpm for 10 min. A mixture of the supernatant and D<sub>2</sub>O was analysed by  $^{19}\text{F}\{^1\text{H}\}$  NMR spectroscopy.  $^{19}\text{F}\{^1\text{H}\}$  NMR (470.40 MHz, H<sub>2</sub>O+D<sub>2</sub>O)  $\delta_{\text{F}}$  -231.0 (5-FDRP, **18**), -231.1 (5'-FDI, **153**), -231.3 (5'-FDA, **17**).

### 5.7.5 Incubations in whole cell cultures of *S. sp.* MA37

#### 5.7.5.1 Incubation of 5'-FDA (17). Day 0

*S. sp.* MA37 (200  $\mu$ L of a glycerol stock) was inoculated into a flask containing ISP Medium N° 2 (50 mL). 5'-FDA (**17**, 5 mg) was added to the flask, and the culture was left to incubate at 30 °C and 180 rpm for 20 days. The culture was then sonicated, and the cell debris removed by centrifugation (8000 rpm, 20 min). The supernatant was freeze-dried and analysed by  $^{19}\text{F}\{^1\text{H}\}$  NMR spectroscopy.  $^{19}\text{F}\{^1\text{H}\}$  NMR (470.40 MHz,  $\text{D}_2\text{O}$ )  $\delta_{\text{F}}$  -216.9 (FAc, **1**), -228.5 (unknown), -228.5 (5-FDR, **26**), -230.9 (unknown), -231.2 (4-FT, **4**), -232.7 (unknown), -233.2 (5-FHPA, **2**).

#### 5.7.5.2 Incubation of 5'-FDA (17). Day 8

*S. sp.* MA37 (200  $\mu$ L of a glycerol stock) was inoculated into a flask containing ISP Medium N° 2 (50 mL), and the culture grown for 8 days at 30 °C and 180 rpm. 5'-FDA **17** (5 mg) was added to the flask, and the culture was left to incubate at 30 °C and 180 rpm for further 12 days. The culture was then sonicated, and the cell debris removed by centrifugation (8000 rpm, 20 min). The supernatant was freeze-dried and analysed by  $^{19}\text{F}\{^1\text{H}\}$  NMR spectroscopy.  $^{19}\text{F}\{^1\text{H}\}$  NMR (470.40 MHz,  $\text{D}_2\text{O}$ )  $\delta_{\text{F}}$  -217.0 (FAc, **1**), -230.8 (5'-FDA, **17**), -231.2 (4-FT, **4**).

#### 5.7.5.3 Incubation of [5',5'- $^2\text{H}_2$ ]-5'-FDA (17a). Day 0

*S. sp.* MA37 (from a seed culture grown for 2 days) was inoculated into a flask containing ISP Medium N° 2 (50 mL). [5',5'- $^2\text{H}_2$ ]-5'-FDA **17a** (8 mg) was added to the flask, and the culture was left to incubate at 30 °C and 180 rpm for 12 days. The culture was then sonicated, and the cell debris removed by centrifugation (8000 rpm, 20 min). The supernatant was freeze-dried and analysed by  $^{19}\text{F}\{^1\text{H}\}$  NMR spectroscopy.  $^{19}\text{F}\{^1\text{H}\}$  NMR (470.40 MHz,  $\text{D}_2\text{O}$ )  $\delta_{\text{F}}$  -216.9 (s, FAc, **1**), -217.5 (t,  $J$  = 8.1 Hz,  $[\text{}^2\text{H}_1]$ -FAc, **1b**), -218.0 (p,  $J$  = 14.0, 8.1 Hz,  $[\text{}^2\text{H}_2]$ -FAc, **1a**), -227.6 (s, unknown), -227.8 (s, unknown), -228.9 (m, unknown, possible 5-FDR, **26**), -230.7 (s, unknown), -231.3 (m, unknown), -231.4 (s, 4-FT, **4**), -231.9 (s, unknown), -232.0 (m,  $[\text{}^2\text{H}_1]$ -4-FT, **4b**), -232.0 (m, unknown), -232.6 (p,  $J$  = 13.0, 5.9 Hz,  $[\text{}^2\text{H}_2]$ -4-FT, **4a**), -233.8 (m,  $[\text{}^2\text{H}_1]$ -5-FHPA), -234.3 (m,  $[\text{}^2\text{H}_2]$ -5-FHPA), -236.1 (m, unknown).



#### **5.7.5.4 Incubation of [5',5'-<sup>2</sup>H<sub>2</sub>]-5'-FDA (17a). Day 4, 5 and 6; without KF**

*S. sp.* MA37 (from a seed culture grown for 2 days) was inoculated into a flask containing ISP Medium N° 2 (50 mL), and the culture grown for 4 days at 30 °C and 180 rpm. [5',5'-<sup>2</sup>H<sub>2</sub>]-5'-FDA **17a** (2 mg, each, approx.) was added on days 4, 5 and 6, and the culture was left to incubate at 30 °C and 180 rpm for further 6 days. The culture was then sonicated, and the cell debris removed by centrifugation (8000 rpm, 20 min). The supernatant was freeze-dried and analysed by <sup>19</sup>F{<sup>1</sup>H} NMR spectroscopy. <sup>19</sup>F{<sup>1</sup>H} NMR (470.40 MHz, D<sub>2</sub>O) δ<sub>F</sub> -217.0 (s, FAc, **1**), -217.5 (t, *J* = 8.1 Hz, [<sup>2</sup>H<sub>1</sub>]-FAc, **1b**), -218.2 (p, *J* = 14.5, 8.1 Hz, [<sup>2</sup>H<sub>2</sub>]-FAc, **1a**), -227.7 (s, unknown), -227.9 (m, unknown), -229.0 (m, unknown), -230.8 (s, unknown), -231.4 (m, unknown), -231.5 (s, 4-FT, **4**), -231.5 (m, unknown), -232.0 (s, unknown), -232.1 (m, [<sup>2</sup>H<sub>1</sub>]-4-FT, **4b**), -232.1 (m, unknown), -232.7 (m, [<sup>2</sup>H<sub>2</sub>]-4-FT, **4a**).

#### **5.7.5.5 Incubation of [5',5'-<sup>2</sup>H<sub>2</sub>]-5'-FDA (17a). Days 4, 5 and 6; with KF**

*S. sp.* MA37 (from a seed culture grown for 2 days) was inoculated into a flask containing ISP Medium N° 2 (50 mL), along with KF (0.2 mM) and the culture grown for 4 days at 30 °C and 180 rpm. [5',5'-<sup>2</sup>H<sub>2</sub>]-5'-FDA **17a** (2 mg, each, approx.) was added on days 4, 5 and 6, and the culture was left to incubate at 30 °C and 180 rpm for further 6 days. The culture was then sonicated, and the cell debris removed by centrifugation (8000 rpm, 20 min). The supernatant was freeze-dried and analysed by <sup>19</sup>F{<sup>1</sup>H} NMR spectroscopy. <sup>19</sup>F{<sup>1</sup>H} NMR (470.40 MHz, D<sub>2</sub>O) δ<sub>F</sub> -216.9 (s, FAc, **1**), -217.4 (t, *J* = 8.1 Hz, [<sup>2</sup>H<sub>1</sub>]-FAc, **1b**), -218.2 (p, *J* = 14.5, 8.1 Hz, [<sup>2</sup>H<sub>2</sub>]-FAc, **1a**), -227.3 (s, unknown), -227.6 (s, unknown), -227.7 (s, unknown), -227.9 (m, unknown), -228.2 (m, unknown), -228.3 (m, unknown), -228.4 (m, unknown), -228.6 (m, unknown), -228.7 (s, unknown), -228.8 (m, unknown), -228.9 (m, unknown), -229.6 (s, unknown), -229.9 (m, unknown), -230.6 (s, unknown), -230.8 (m, unknown), -231.0 (s, unknown), -231.2 (m, unknown), -231.2 (s, 4-FT, **4**), -231.4 (m, unknown), -231.8 (s, unknown), -231.9 (m, [<sup>2</sup>H<sub>1</sub>]-4-FT, **4b**), -231.9 (m, unknown), -232.5 (m, [<sup>2</sup>H<sub>2</sub>]-4-FT, **4a**), -232.6 (s, unknown), -233.1 (s, 5-FHPA, **2**), -234.8 (m, unknown).

## 5.8 References

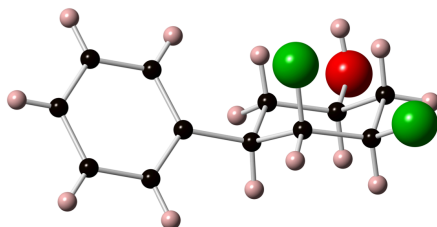
- 
- <sup>1</sup> W. L. F. Armarego, C. L. L. Chai; *Purification of Laboratory Chemicals*, 5<sup>th</sup> Ed. Butterworth-Heinemann (Elsevier): Oxford, **2003**.
- <sup>2</sup> L. Ma, A. Bartholomé, M. H. Tong, Z. Qin, Y. Yu, T. Shepherd, K. Kyeremeh, H. Deng, D. O'Hagan; *Chem. Sci.* **2015**, 6, 1414.
- <sup>3</sup> S. Dall'Angelo, N. Bandaranayaka, A. D. Windhorst, D. J. Vugts, D. Van der Born, M. Onega, L. F. Schweiger, M. Zanda, D. O'Hagan; *Nuc. Med. Biol.* **2013**, 40, 464.
- <sup>4</sup> D. O'Hagan, J. H. Naismith, C. Schaffrath, C. Dong, J. B. Spencer, F. Huang; *PTC Int. Appl.* **2004**, WO2004078914.
- <sup>5</sup> K. Graf, C. L. Rühl, M. Rudolph, F. Rominger, A. S. K. Hashmi, *Angew. Chem.* **2013**, 52, 12727 – 12731.
- <sup>6</sup> M. Onega, PhD Thesis (University of St Andrews), **2009**.
- <sup>7</sup> W. Li, X. Yin, S. W. Schneller; *Bioorg. Med. Chem. Lett.* **2008**, 18, 220.
- <sup>8</sup> X.-G. Li, S. Dall'Angelo, L. F. Schweiger, M. Zanda, D. O'Hagan; *Chem. Commun.* **2012**, 48, 5247.
- <sup>9</sup> C. L. Coombes, C. J. Moody; *J. Org. Chem.* **2008**, 73, 6758.
- <sup>10</sup> B. Linclau, Z. Wang, G. Compain, V. Paumelle, C. Q. Fontenelle, N. Wells, A. Weymouth-Wilson; *Angew. Chem. Int. Ed.* **2016**, 55, 674.
- <sup>11</sup> J. Sangster, *J. Phys. Chem. Ref. Data*, **1989**, 18, 1111-1227.
- <sup>12</sup> C. Hansch, A. Leo, D. Hoekman, *Exploring QSAR – Hydrophobic, Electronic, and Steric Constants*, Washington, DC: American Chemical Society, **1995**.
- <sup>13</sup> P. C. von der Ohe, R. Kühne, R.-U. Ebert, R. Altenburger, M. Liess, G. Schüürmann, *Chem. Res. Toxicol.* **2005**, 18, 536 – 555.
- <sup>14</sup> S. E. Ruzin; *Plant Microtechnique and Microscopy*, Oxford University Press, **1999**.

---

<sup>15</sup> A. Vandemeulebroucke, S. De Vos, E. Van Holsbeke, J. Steyaert, W. Versées; *J. Biol. Chem.* **2008**, 283, 22272.

## Appendix – Crystallographic Data

### 3,4-Difluoro-5-phenylcyclohexan-1-ol (80)



**Code:** agdh1

**Structure Type:** Crystal

**Chemical Formula:** C<sub>12</sub> H<sub>14</sub> F<sub>2</sub> O

**Display Formula:** C<sub>12</sub> H<sub>14</sub> F<sub>2</sub> O

**Spacegroup:** *P* 2<sub>1</sub> 2<sub>1</sub> 2<sub>1</sub>

*(Allows Chirality)*

**Crystal System:** Orthorhombic

a: 5.2589 Å

b: 9.7524 Å

c: 20.4698 Å

**Asymmetric Unit:** 29 sites

**Unit Cell:** 116 sites per unit cell

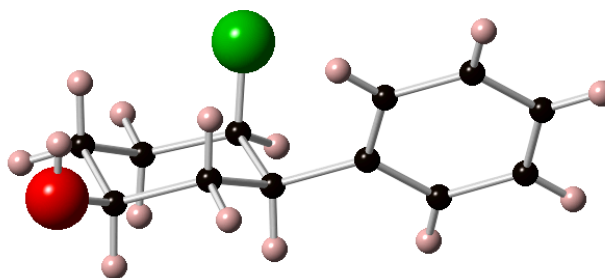
**Site Density:** 0.1105 sites/Å<sup>3</sup>

**Visible Atoms:** 29

**Cell Volume:** 1049.832 Å<sup>3</sup>

**Density:** 1.3428 g/cm<sup>3</sup>

**4-Fluoro-3-phenylcyclohexan-1-ol (81)**



**Code:** agdh6

**Structure Type:** Crystal

**Chemical Formula:**  $C_{12}H_{14.8}FO$

**Display Formula:**  $C_{12}H_{14.8}FO$

**Spacegroup:**  $P 2_1 2_1 2_1$

*(Allows Chirality)*

**Crystal System:** Orthorhombic

**a:** 5.3302 Å

**b:** 9.6528 Å

**c:** 20.0332 Å

**Asymmetric Unit:** 30 sites

**Unit Cell:** 120 sites per unit cell

**Site Density:** 0.1164 sites/Å<sup>3</sup>

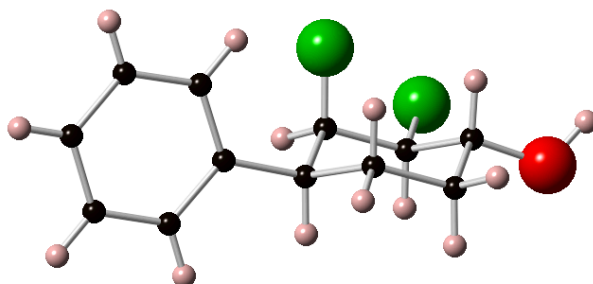
0.1118 atoms/Å<sup>3</sup>

**Visible Atoms:** 30

**Cell Volume:** 1030.743 Å<sup>3</sup>

**Density:** 1.2504 g/cm<sup>3</sup>

**2,3-Difluoro-4-phenylcyclohexan-1-ol (82)**



**Code:** agdh2

**Structure Type:** Crystal

**Chemical Formula:** C<sub>12</sub> H<sub>14</sub> F<sub>2</sub> O

**Display Formula:** C<sub>12</sub> H<sub>14</sub> F<sub>2</sub> O

**Spacegroup:** *P* 2<sub>1</sub>/*a*

**Crystal System:** Monoclinic

a: 13.5173 Å

b: 5.5730 Å

c: 27.4909 Å

β: 90.469°

**Asymmetric Unit:** 58 sites

**Unit Cell:** 232 sites per unit cell

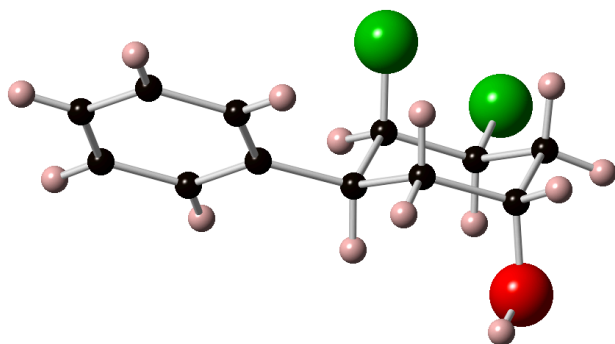
**Site Density:** 0.1120 sites/Å<sup>3</sup>

**Visible Atoms:** 29

**Cell Volume:** 2070.884 Å<sup>3</sup>

**Density:** 1.3614 g/cm<sup>3</sup>

**3,4-Difluoro-5-phenylcyclohexan-1-ol (83)**



**Code:** agdh4

**Structure Type:** Crystal

**Chemical Formula:** C<sub>12</sub> H<sub>14</sub> F<sub>2</sub> O

**Display Formula:** C<sub>12</sub> H<sub>14</sub> F<sub>2</sub> O

**Spacegroup:** *P* 2<sub>1</sub>/*c*

**Crystal System:** Monoclinic

a: 10.5360 Å

b: 19.9830 Å

c: 10.2720 Å

β: 107.972°

**Asymmetric Unit:** 58 sites

**Unit Cell:** 232 sites per unit cell

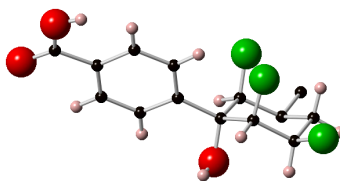
**Site Density:** 0.1128 sites/Å<sup>3</sup>

**Visible Atoms:** 29

**Cell Volume:** 2057.154 Å<sup>3</sup>

**Density:** 1.3705 g/cm<sup>3</sup>

**4-(2,3,6-trifluoro-1-hydroxycyclohexyl)benzoic acid (90)**



**Code:** agdh5

**Structure Type:** Crystal

**Chemical Formula:**  $C_{41.5} H_{34} F_9 O_9$

**Display Formula:**  $C_{14} H_{11} F_3 O_3$

**Spacegroup:**  $P 1$

*(Allows Chirality; Polar)*

**Crystal System:** Triclinic

a: 6.0731 Å

b: 13.2193 Å

c: 23.0241 Å

$\alpha$ : 94.519°

$\beta$ : 93.658°

$\gamma$ : 101.657°

**Asymmetric Unit:** 187 sites

**Unit Cell:** 187 sites per unit cell

**Site Density:** 0.1040 sites/Å<sup>3</sup>

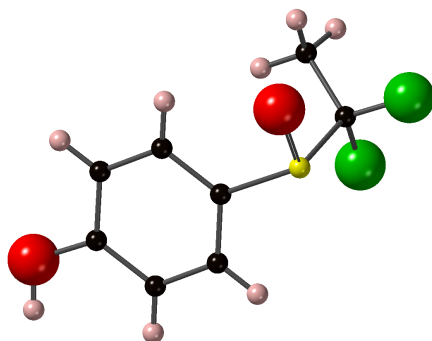
**Visible Atoms:** 31

**Cell Volume:** 1798.637 Å<sup>3</sup>

**Density:** 1.5652 g/cm<sup>3</sup>



**4-((1,1-Difluoroethyl)sulfinyl)phenol (113)**



**Code:** agdh14

**Structure Type:** Crystal

**Chemical Formula:** C<sub>8</sub> H<sub>8</sub> F<sub>2</sub> O<sub>2</sub> S

**Display Formula:** C<sub>8</sub> H<sub>8</sub> F<sub>2</sub> O<sub>2</sub> S

**Spacegroup:** *P* 2<sub>1</sub> 2<sub>1</sub> 2<sub>1</sub>

*(Allows Chirality)*

**Crystal System:** Orthorhombic

a: 8.5526 Å

b: 9.2693 Å

c: 10.9710 Å

**Asymmetric Unit:** 21 sites

**Unit Cell:** 84 sites per unit cell

**Site Density:** 0.0966 sites/Å<sup>3</sup>

**Visible Atoms:** 21

**Cell Volume:** 869.744 Å<sup>3</sup>

**Density:** 1.5748 g/cm<sup>3</sup>

## Appendix – Publications

A. Rodil, S. Bosisio, M. S. Ayoub, L. Quinn, D. B. Cordes, A. M. Z. Slawin, C. D. Murphy, J. Michel, D. O'Hagan; *Chem. Sci.* **2018**, *9*, 3023.

R. Tomita, N. Al-Maharik, A. Rodil, M. Bühl, D. O'Hagan; *Org. Biomol. Chem.* **2018**, *16*, 1113.

Owing to licensing issues, the full text of the second article is not included in the electronic version of this thesis. A copy of the accepted manuscript for this article can be found in the St Andrews Research Repository at <http://hdl.handle.net/10023/16826>.

Cite this: *Chem. Sci.*, 2018, 9, 3023

# Metabolism and hydrophilicity of the polarised 'Janus face' all-*cis* tetrafluorocyclohexyl ring, a candidate motif for drug discovery†

Andrea Rodil,<sup>a</sup> Stefano Bosisio,<sup>b</sup> Mohammed Salah Ayoup,<sup>ad</sup> Laura Quinn,<sup>c</sup> David B. Cordes,<sup>id</sup> <sup>a</sup> Alexandra M. Z. Slawin,<sup>id</sup> <sup>a</sup> Cormac D. Murphy,<sup>c</sup> Julien Michel<sup>\*b</sup> and David O'Hagan<sup>id</sup> <sup>\*a</sup>

The metabolism and polarity of the all-*cis* tetra-fluorocyclohexane motif is explored in the context of its potential as a motif for inclusion in drug discovery programmes. Biotransformations of phenyl all-*cis* tetra-, tri- and di- fluoro cyclohexanes with the human metabolism model organism *Cunninghamella elegans* illustrates various hydroxylated products, but limited to benzylic hydroxylation for the phenyl all-*cis* tetrafluorocyclohexyl ring system. Evaluation of the lipophilicities (log *P*) indicates a significant and progressive increase in polarity with increasing fluorination on the cyclohexane ring system. Molecular dynamics simulations indicate that water associates much more closely with the hydrogen face of these Janus face cyclohexyl rings than the fluorine face owing to enhanced hydrogen bonding interactions with the polarised hydrogens and water.

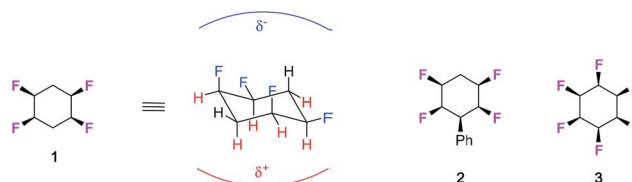
Received 18th January 2018  
Accepted 17th February 2018

DOI: 10.1039/c8sc00299a

rsc.li/chemical-science

## Introduction

Approximately one third of drugs on the market or in development contain at least one fluorine atom<sup>1</sup> and around a third of herbicides historically,<sup>2</sup> and half the commercial pesticides introduced in the last period (2010–2016)<sup>3</sup> contain fluorine atoms. The element is also important in organic materials with applications in next generation displays<sup>4</sup> and high value materials.<sup>5</sup> The investigation of new products bearing fluorinated moieties is an ever expanding field, given the particular properties that fluorine bestows on organic compounds.<sup>6</sup> We have recently synthesised all-*cis* 1,2,4,5-tetrafluorocyclohexane ring systems such as **1** and **2** as a novel motif in organic chemistry.<sup>7</sup> This tetrafluorocyclohexane isomer displays a particular polar property across the cyclohexane, largely because all of the fluorines are on one face of the ring and there are two 1,3-diaxial C–F bonds, with dipoles orientated parallel to each other.



As an extreme example, we extended this concept to the preparation and analysis of all-*cis* hexafluorocyclohexane **3** which is even more polar due to a cyclohexane ring accommodating three axial C–F bonds.<sup>8a</sup> This cyclohexane has been referred to as a 'Janus'-like molecule,<sup>8b</sup> because of its very well differentiated faces, and recent experimental and theoretical studies have indicated that these rings will coordinate cations to the fluorine face and anions to the hydrogen face, consistent with the electrostatic polarity of the ring system.<sup>9</sup> The conformational and polar properties of these multi-vicinal fluorinated aliphatics is beginning to attract the attention of the synthesis community and new methods are emerging for their preparation, for example, from the Gilmour,<sup>9</sup> Jacobsen<sup>10</sup> and Carreira<sup>11</sup> laboratories. For the cyclohexanes, a recent report from Glorius's laboratory<sup>12</sup> has demonstrated the direct catalytic hydrogenation of fluorinated aromatics to generate all-*cis* fluorinated cyclohexanes in a single step, and this methodology promises to make compounds such as cyclohexanes **1** and **3** much more accessible to the organic chemistry community. With these developments in synthesis methods, we believe the cyclohexane motif merits exploration as a candidate substituent for agrochemicals or pharmaceutical drug discovery programmes.

<sup>a</sup>EaStChem School of Chemistry, University of St Andrews, North Haugh, St Andrews, Fife KY16 9ST, UK. E-mail: do1@st-andrews.ac.uk

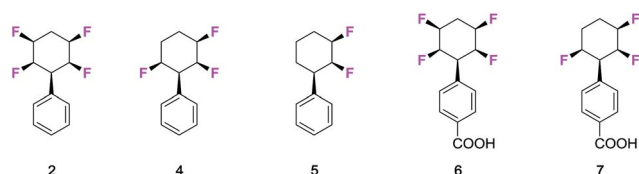
<sup>b</sup>EaStChem School of Chemistry, University of Edinburgh, Joseph Black Building, David Brewster Road, Edinburgh, EH9 3FJ, UK. E-mail: julien.michel@ed.ac.uk

<sup>c</sup>UCD School of Biomolecular and Biomedical Sciences, University College Dublin, Belfield, Dublin, Ireland. E-mail: cormac.d.murphy@ucd.ie

<sup>d</sup>Department of Chemistry, Faculty of Science, Alexandria University, P.B. 426 Ibrahimia, Egypt

† Electronic supplementary information (ESI) available. CCDC 1817386–1817390. For ESI and crystallographic data in CIF or other electronic format see DOI: 10.1039/c8sc00299a

Immediate questions which arise are how will these selectively fluorinated cyclohexane rings be metabolised and how lipophilic are these ring systems. It is commonly understood that increasing the level of fluorination of an organic motif will generally result in increasing its resistance to metabolism at certain sites.<sup>13</sup> Also, the prevailing dogma is that increased levels of fluorination render a motif more lipophilic and, thus, its introduction would have a tendency to raise log *P* values in a manner detrimental to judicious selection in medicinal chemistry. However, it is more complex than that, and Müller and Carreira have exemplified this extensively in recent contributions *e.g.* mapping log *P*s of RCH<sub>3</sub> compounds through progressive fluorination to RCF<sub>3</sub>, where intermediate fluorinations (RCH<sub>2</sub>F & RCF<sub>2</sub>H) decrease lipophilicity.<sup>14</sup> It is a feature of these ring systems,<sup>15</sup> where the fluorines have a relative stereochemistry such that they are all on one face of the cyclohexane, that the rings become polar, and thus increasing fluorination could reasonably increase hydrophilicity. Thus we set out to explore the nature of these ring systems in the context of their properties and potential as a novel motif for inclusion in bioactive research programmes. To that end we focus on phenylcyclohexane **2**, because it is readily prepared<sup>7c</sup> and has been shown to be amenable to a range of synthetic transformations and diversification.<sup>16,7b</sup> The study compared the metabolism of **2** to close analogues **4–7** by incubation with the human metabolism model organism *Cunninghamella elegans*.<sup>17</sup> Lipophilicity trends (log *P*) were also explored comparing cyclohexanes with four, three, two and no fluorine atoms. Lastly, a molecular dynamics simulation study was carried out to elucidate the structural basis of the observed lipophilicity trends.



## Results and discussion

### Biotransformations with *Cunninghamella elegans*

The fungus *Cunninghamella elegans* represents a well-established model for drug metabolism in mammals due to its ability to biotransform and degrade a wide range of xenobiotics.<sup>18</sup> The organism contains a range of cytochrome P<sub>450</sub> enzymes and this gives an oxidative metabolic profile which mimics phase-I oxidative metabolism. In order to investigate how the tetrafluorocyclohexyl motif may be metabolised, we have explored the incubation of phenyl tetrafluorocyclohexanes and also compounds with three and two fluorine atoms. Five compounds were investigated in total, three of which were the phenyl derivatives **2**, **4** and **5**, and two were the benzoic acid derivatives **6** and **7**.<sup>7d</sup> Incubations with *C. elegans* were carried out in triplicate in submerged liquid cultures. In each case, the incubations were worked up after three days and products were extracted and analysed.

Phenylcyclohexane **2** gave rise to only one obvious metabolite in a conversion of around 30%. This product arose by direct hydroxylation at the benzylic position of **2** to give benzyl alcohol **8**. Only one product as a single isomer could be detected, with the hydroxyl group configured *anti* to the adjacent fluorine atoms of the cyclohexane ring. The identity of **8** and its stereochemistry was confirmed by X-ray structure analysis.

Phenyl trifluorocyclohexane **4** was similarly incubated with the fungus and it too gave rise to the analogous benzyl hydroxylated product **9**. The extent of microbial conversion was approximately 50% after the three day incubation. The residual **4** was assayed for enantiomeric purity by chiral HPLC, and it was shown to be almost racemic, thus there is no indication that the microbial hydroxylation was significantly enantioselective. Finally in this series, difluorocyclohexane **5** was subject to a similar incubation with *C. elegans*. This compound was completely and extensively metabolised, and it generated a much greater product profile of which compounds **10–13** were isolated. Compounds **11–13** were characterised by X-ray crystallography as illustrated in Fig. 1. Monohydroxylated products

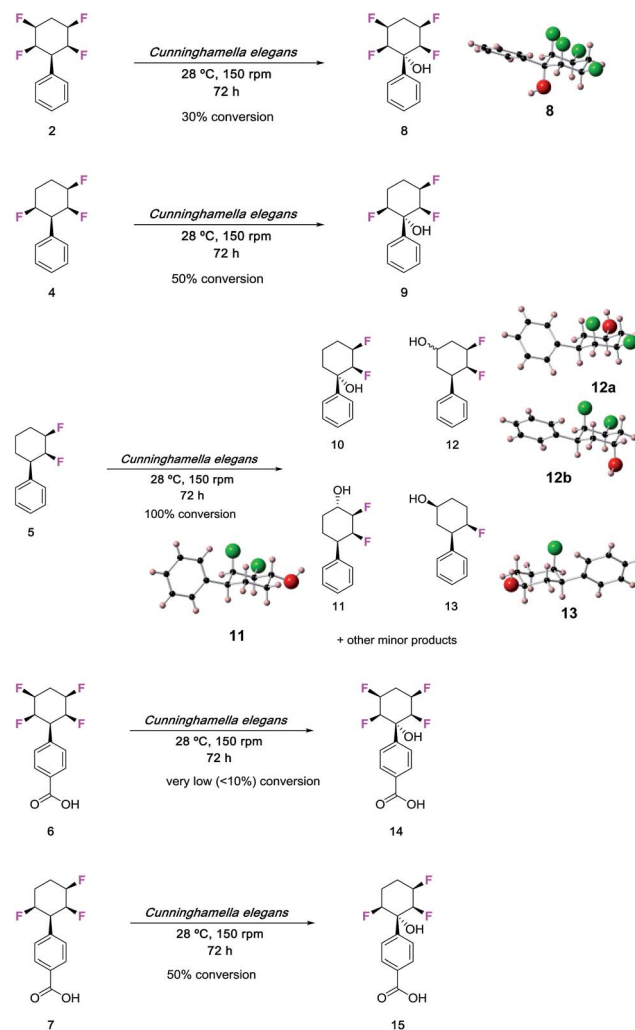


Fig. 1 Biotransformations of selectively fluorinated phenyl fluorocyclohexanes **2** and **4–7** by *C. elegans*. Some of the products were crystalline and amenable to X-ray structure analysis.



10–12, can be rationalised by direct methylene  $P_{450}$  type hydroxylations, however the monofluorinated cyclohexanol 13 is less easily rationalised and presumably arises from a series of biotransformations involving fluoride elimination. More generally, it is clear that removal of two of the ring fluorines from positions 2 and 3 of the phenyl all-*cis* tetracyclohexyl ring system has rendered the aliphatic ring much more susceptible to metabolism.

The benzoic acids 6 and 7 were also incubated with *C. elegans*. The tetrafluorocyclohexyl benzoic acid 6 was poorly biotransformed and only a very low conversion to alcohol 14 was obvious after the three day incubation. Trifluorocyclohexyl benzoic acid 7 was more readily transformed, but only to benzylalcohol 15 (~50% conversion). This product was isolated and crystallised and X-ray analysis confirmed its structure. Again, in order to explore any enantioselectivity for this biotransformation, the methyl ester of the residual carboxylic acid 7 was analysed by chiral HPLC and this indicated a very low enantioselectivity, thus in a similar outcome to substrate 4, there was no obvious selectivity for 7 by the hydroxylation enzyme involved.

### Lipophilicity study of selectively fluorinated phenyl cyclohexanes

An important measure of the druggability of a substituent is its lipophilicity,<sup>19</sup> and given the polarity of the phenyl all-*cis* tetrafluorocyclohexyl moiety it was of interest to explore the relative log *P*s of various analogous compounds. log *P*'s were measured by reverse phase HPLC (AcCN 60%: water 40%, with TFA 0.05%), as previously described.<sup>20</sup> The measured log *P*'s of a series of phenyl fluorocyclohexane derivatives are summarised in Fig. 2 and against a series of compounds, of known log *P* values, which were re-measured for comparison, including biphenyl 16 and phenylcyclohexane 17.

It is clear that there is a significant reduction in log *P* with increasing fluorination. Phenyldifluorocyclohexane 5 (log *P* 3.30) is significantly more polar than the phenylcyclohexane (log *P* 4.99), and then both the tri- and tetra- fluoro

cyclohexanes progressively increase in polarity (log *P*s of 2.64 and 2.58 respectively) with additional fluorine atoms. An interesting comparison on log *P*s can be made with the two trifluorinated stereoisomers 4 and 18. Compound 4 is more polar, and this presumably arises as it has a preferred diaxial arrangement of the C2 and C6 C–F bonds.<sup>7c</sup> This parallel alignment can be expected to increase the molecular dipole relative to isomer 18 which has one of these fluorines lying in an equatorial orientation.

The study extended to substituted aryls of the benzoic acids 6 and 7 and the anilines 21 and 22 (ref. 7b) as illustrated in Fig. 3. In each case both the trifluoro- and tetrafluoro- cyclohexanes are around two log *P* units more lipophilic than the non-fluorinated cyclohexanes 20 and 24, whereas the phenyl derivatives 19 and 23 lie in between. There is a clear trend that selective fluorinations around the ring increases the polarity of the cyclohexane.

### Computational analysis of lipophilicity trends

Molecular dynamics (MD) simulations were carried out for phenylcyclohexanes 2, 4, 5, 16 and 17 to clarify the mechanisms by which progressive fluorination decreases lipophilicity. log *P*s were predicted by computing absolute solvation free energies in aqueous and organic phases using explicit solvent molecular dynamics simulations.<sup>21</sup> Fig. 4 shows a comparison of calculated (log  $P_{\text{pred}}$ ) and measured (log  $P_{\text{exp}}$ ) log *P* values, as well as calculated solvation free energies in aqueous ( $\Delta G_{\text{aq}}$ ) and cyclohexane ( $\Delta G_{\text{org}}$ ) phases. Overall, the log *P* calculations are in good agreement with the experimental data (Kendall tau  $0.5 \pm 0.1$  and mean unsigned error  $0.77 \pm 0.07$  log *P* units). Inspection of the solvation free energies shows that the trend for decreased log *P* upon increased fluorination is due to a more rapid decrease in solvation free energies in the aqueous phase (from ca.  $-2.5$  to  $-5.2$  kcal mol<sup>-1</sup> for 16 and 2 respectively) vs. the cyclohexane phase (ca.  $-7.5$  kcal mol<sup>-1</sup> for all compounds).

Further insights were investigated to help rationalise the calculated differences in hydration free energies by grid-cell theory (GCT) analyses of the MD simulation trajectories.<sup>22</sup>

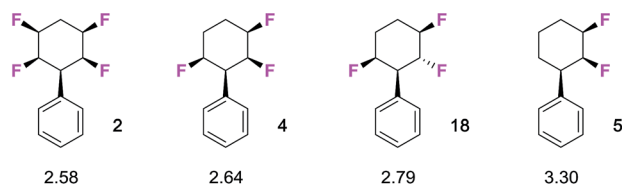
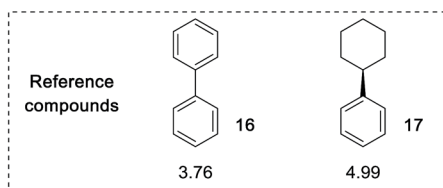


Fig. 2 Measured<sup>20</sup> log *P* values for compounds selectively fluorinated phenylcyclohexanes and reference compounds 16 and 17. Increasing fluorination lowers log *P* consistent with increasing hydrophilicity.

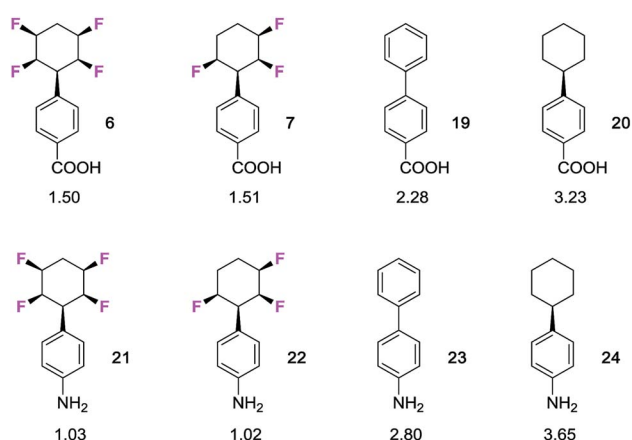


Fig. 3 Measured<sup>20</sup> log *P* values for benzoic acid and aniline derivatives of selectively fluorinated cyclohexanes.

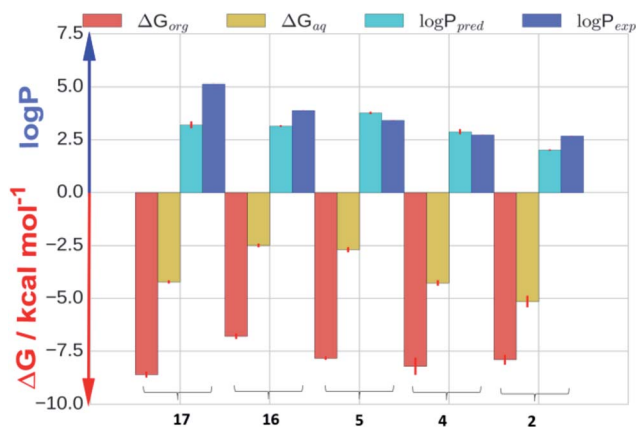


Fig. 4 The positive y-axis depicts a comparison between calculated (cyan) and measured (blue)  $\log P$  values for compounds 2, 4, 5, 16 and 17. The negative y-axis depicts calculated solvation free energy in cyclohexane,  $\Delta G_{org}$  (red), and aqueous,  $\Delta G_{aq}$  (yellow), phases.

GCT is a MD trajectory post-processing method that spatially resolves the water contribution to enthalpies, entropies and free energies of the hydration for small molecules, host/guests and protein-ligands complexes.<sup>23</sup>

Fig. 5 depicts spatially resolved hydration thermodynamics around the non-fluorinated cyclohexane 17 and the tetrafluorinated cyclohexane 2. Comparison of water density contours show water structuring above and below the  $\pi$ -cloud of

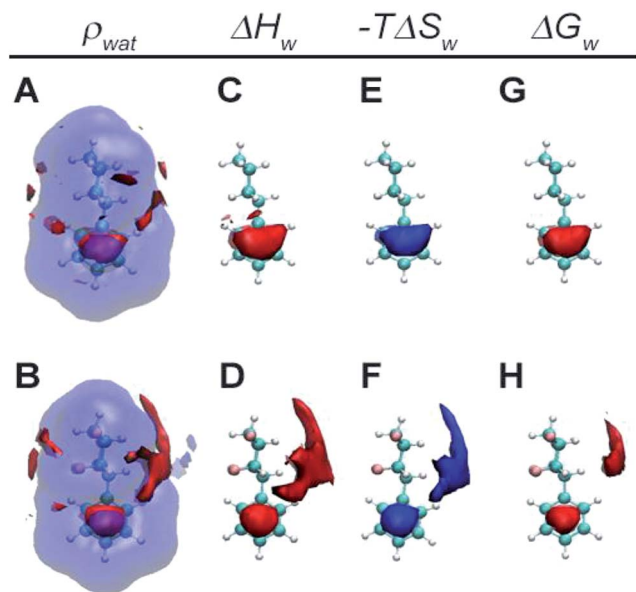


Fig. 5 Spatial resolution of hydration thermodynamics around 17 and 2. Panels A and B show isocontours for density (red:  $\rho_{wat} > 2.33$  bulk density, blue:  $\rho_{wat} < 0.5$  bulk density). Panels C and D show isocontours for regions where water is enthalpically stabilised with respect to bulk water (red:  $\Delta H_w < -0.0055$  kcal mol<sup>-1</sup> A<sup>-3</sup>). Panels E and F show isocontours for regions where water is entropically destabilised with respect to bulk water (blue:  $-T\Delta S_w > 0.0033$  kcal mol<sup>-1</sup> A<sup>-3</sup>). Panels G and H show isocontours for regions where water is more stable than bulk water (red:  $\Delta G_w < -0.0055$  kcal mol<sup>-1</sup> A<sup>-3</sup>).

the phenyl ring due to the expected weak hydrogen bonding interactions in this region. In addition the four fluorine atoms in 2 induce further structuring of water around the cyclohexyl moiety, with a more pronounced effect around the hydrogen face of the cyclohexane (panels A and B). Owing to the different polarities of the cyclohexyl ring in 2, water near the fluorine-face preferentially orients hydrogen atoms towards the ring, whereas water near the hydrogen-face preferentially orients oxygen atoms towards the ring. Water near the hydrogen-face is more enthalpically stabilised and entropically destabilised with respect to bulk, whereas the energetics are not significantly different from the bulk in the vicinity of the fluorine face (panels C and D and E and F). Overall favourable enthalpic contributions offset unfavourable entropic contributions for water near the hydrogen face and water in this region makes additional favourable contributions to the hydration free energy (panels G and H). Therefore the decreased lipophilicity of 2 with respect to 17 is attributed to enhanced hydrogen bonding interactions between water and the hydrogen face of the all-*cis* tetrafluorocyclohexane ring.

## Conclusions

The all-*cis* tetrafluorocyclohexane motif has been recognised to have particularly polar properties and the ease of synthesis of the phenyl derivative 2 has prompted us to investigate its properties further as it emerges as a building block for the introduction of this new motif into medicinal chemistry and other bioactives discovery programmes. The metabolism of the phenyl cyclohexane derivatives 2, 4–7 with varying levels of fluorination was explored in incubations with *Cunninghamella elegans*. This fungus has been used as a microbial model for mammalian metabolism. In the present study we observed that increasing the degree of fluorination of cyclohexyl ring leads to a more stable xenobiotic. The phenyl all-*cis* tetrafluorocyclohexane 2 was significantly less metabolised than the trifluoro-4 and then difluoro-5, the latter of which was extensively metabolised. In the case of 2, 4, 6 and 7 metabolism is confined to benzylic hydroxylation.

A systematic  $\log P$  evaluation of these ring systems shows an increase in hydrophilicity with increasing fluorination, and for the phenyl all-*cis* tetrafluorocyclohexanes (including anilines and benzoic acids) there is a maximal effect. These ring systems are at least two full  $\log P$  units (100 fold) more hydrophilic than their non-fluorinated cyclohexane counterparts.

Molecular dynamics simulations reproduce the experimental trends and suggest that the decreased lipophilicity of 2 is due to enhanced hydrogen bonding interactions of water molecules with the hydrogen face of the cyclohexane ring with respect to bulk water. The orientation of the water near this face of the ring was consistent with the hydrogen bonding donor ability of the polarised hydrogens of the ring.

This contrasts with the energetics of water near the fluorine face of the ring which are comparable to bulk water. Altogether these studies indicate that metabolism of the all-*cis* tetrafluorocyclohexyl motif is slow, and that the ring system is significantly hydrophilic for an aliphatic motif. These factors





add to the unique facially polarised aspect of this motif and make it an attractive option for inclusion in medicinal chemistry or crop protection studies.

## Conflicts of interest

There are no conflicts of interest.

## Acknowledgements

This work was supported by the Initial Training Network, FLUOR21, funded by the FP7 Marie Curie Actions of the European Commission (FP7-PEOPLE-2013-ITN-607787). JM is supported by a University Research Fellowship from the Royal Society. The research leading to these results has received funding from the European Research Council under the EU 7th Framework Programme (FP7/2007-2013)/ERC No 336289. We acknowledge the EPSRC UK National Mass Spectrometry Facility at Swansea University. We also thank Drs Guillaume Berthon, Adam Burris and Andreas Beck of Syngenta Crop Protection AG, Stein Switzerland for assistance in log *P* determinations.

## Notes and references

- (a) J. Wang, M. Sanchez-Rosello, J. L. Acena, C. del Pozo, A. E. Sorochinsky, S. Fustero, V. A. Soloshonok and H. Liu, *Chem. Rev.*, 2014, **114**, 2432–2506; (b) D. O'Hagan and D., *J. Fluorine Chem.*, 2010, **131**, 1071–1081; (c) S. Purser, P. R. Moore, S. Swallow and V. Gouverneur, *Chem. Soc. Rev.*, 2008, **37**, 320–330; (d) J.-P. Begue and D. Bonnet-Delpon, *J. Fluorine Chem.*, 2006, **127**, 992–1012; (e) K. Müller, C. Faeh and F. Diederich, *Science*, 2007, **317**, 1881–1886; (f) W. K. Hagmann, *J. Med. Chem.*, 2008, **51**, 4359–4369; (g) K. L. Kirk, *J. Fluorine Chem.*, 2006, **127**, 1013–1029.
- T. Fujiwara and D. O'Hagan, *J. Fluorine Chem.*, 2014, **167**, 16–29.
- P. Jeschke, *Pest Manage. Sci.*, 2017, **73**, 1053–1066.
- (a) M. Alaasar, M. Prehm, S. Poppe and C. Tschierske, *Chem.-Eur. J.*, 2017, **23**, 5541–5556; (b) P. Kirsch, *J. Fluorine Chem.*, 2015, **177**, 29–36; (c) N. Al-Maharik, P. Kirsch, A. M. Z. Slawin and D. O'Hagan, *Tetrahedron*, 2014, **70**, 4626–4630; (d) M. Bremer, P. Kirsch, M. Klasen-Memmer and K. Tarumi, *Angew. Chem., Int. Ed.*, 2013, **52**, 8880–8896.
- (a) W. Feng, P. Long, Y. Feng and Y. Li., *Adv. Sci.*, 2016, **3**, 1500413; (b) R. J. Kashtiban, M. A. Dyson, R. R. Nair, R. Zan, S. L. Wong, Q. Ramasse, A. K. Geim, U. Bangert and J. Sloan, *Nat. Commun.*, 2014, **5**, 5545.
- (a) C. Thiehoff, Y. P. Rey and R. Gilmour, *Isr. J. Chem.*, 2017, **57**, 92–100; (b) D. L. Orsi and R. A. Altman, *Chem. Commun.*, 2017, **53**, 7168–7181.
- (a) A. J. Durie, A. M. Z. Slawin, T. Lebl, P. Kirsch and D. O'Hagan, *Chem. Commun.*, 2012, **48**, 9643–9645; (b) A. J. Durie, T. Fujiwara, R. Cormanich, M. Bühl, A. M. Z. Slawin and D. O'Hagan, *Chem.-Eur. J.*, 2014, **20**, 6259–6263; (c) A. J. Durie, T. Fujiwara, N. Al-Maharik, A. M. Z. Slawin and D. O'Hagan, *J. Org. Chem.*, 2014, **79**, 8228–8233; (d) M. S. Ayoup, D. B. Cordes, A. M. Z. Slawin and D. O'Hagan, *Org. Biomol. Chem.*, 2015, **13**, 5621–5624; (e) M. S. Ayoup, D. B. Cordes, A. M. Z. Slawin and D. O'Hagan, *Beilstein J. Org. Chem.*, 2015, **11**, 2671–2676.
- (a) N. S. Keddie, A. M. Z. Slawin, T. Lebl, D. Philp and D. O'Hagan, *Nat. Chem.*, 2015, **7**, 483–488; (b) N. Santschi and R. Gilmour, *Nat. Chem.*, 2015, **7**, 467–468.
- (a) F. Scheidt, P. Selter, N. Santschi, M. C. Holland, D. V. Dudenko, C. Daniliuc, C. Mück-Lichtenfeld, M. R. Hansen and R. Gilmour, *Chem.-Eur. J.*, 2017, **23**, 6142–6149; (b) I. G. Molinár and R. Gilmour, *J. Am. Chem. Soc.*, 2016, **138**, 5004–5007; (c) I. G. Molnár, C. Thiehoff, M. C. Holland and R. Gilmour, *ACS Catal.*, 2016, **6**, 7167–7173.
- S. M. Banik, J. W. Medley and E. N. Jacobsen, *J. Am. Chem. Soc.*, 2016, **138**, 5000–5003.
- S. Fischer, N. Huwyler, S. Wolfrum and E. M. Carreira, *Angew. Chem., Int. Ed.*, 2016, **55**, 2555–2558.
- M. P. Wiesenfeldt, Z. Nairoukh, W. Li and F. Glorius, *Science*, 2017, **357**, 908–912.
- (a) T. Bright, F. Dalton, V. L. Elder, C. D. Murphy, N. K. O'Connor and G. Sandford, *Org. Biomol. Chem.*, 2013, **11**, 1135–1142; (b) M. J. Shaughnessy, A. Harsanyi, J. Li, T. Bright, C. D. Murphy and G. Sandford, *ChemMedChem*, 2014, **9**, 733–736.
- (a) Q. A. Huchet, N. Trapp, B. Kuhn, B. Wagner, H. Fischer, N. A. Kratochwil, E. M. Carreira and K. Müller, *J. Fluorine Chem.*, 2017, **198**, 34–46; (b) R. Vorberg, N. Trapp, D. Zimmerli, B. Wagner, H. Fischer, N. A. Kratochwil, M. Kansy, E. M. Carreira and K. M. Müller, *ChemMedChem*, 2016, **11**, 2216–2239.
- (a) B. E. Ziegler, M. Lecours, R. A. Marta, J. Featherstone, E. Fillion, W. S. Hopkins, V. Steinmetz, N. S. Keddie, D. O'Hagan and T. B. McMahon, *J. Am. Chem. Soc.*, 2016, **138**, 7460–7463; (b) M. J. Lecours, R. A. Marta, V. Steinmetz, N. S. Keddie, E. Fillion, D. O'Hagan, T. B. McMahon and W. S. Hopkins, *J. Phys. Chem. Lett.*, 2017, **8**, 109–113.
- (a) T. Bykova, N. Al-Maharik, A. M. Z. Slawin and D. O'Hagan, *Beilstein J. Org. Chem.*, 2017, **13**, 728–733; (b) T. Bykova, N. Al-Maharik, A. M. Z. Slawin and D. O'Hagan, *Org. Biomol. Chem.*, 2016, **14**, 1117–1123.
- M. Hezari and P. J. Davis, *Drug Metab. Dispos.*, 1993, **21**, 259–267.
- (a) W. Palmer-Brown, B. Dunne, Y. Ortin, M. A. Fox, G. Sandford and C. D. Murphy, *Xenobiotica*, 2017, **47**, 763–770; (b) J. Amadio, E. Casey and C. D. Murphy, *Appl. Microbiol. Biotechnol.*, 2013, **97**, 5955–5963; (c) J. Amadio and C. D. Murphy, *Appl. Microbiol. Biotechnol.*, 2010, **86**, 345–351; (d) L. Quinn, R. Dempsey, E. Casey, A. Kane and C. D. Murphy, *J. Ind. Microbiol. Biotechnol.*, 2015, **42**, 799–806.
- (a) C. A. Lipinski, *Adv. Drug Delivery Rev.*, 2016, **101**, 34–41; (b) C. A. Lipinski, F. Lombardo, B. W. Dominy and P. J. Feeney, *Adv. Drug Deliv. Rev.*, 1997, **23**, 3–25.
- (a) R. Tomita, N. Al-Maharik, A. Rodil, M. Bühl and D. O'Hagan, *Org. Biomol. Chem.*, 2018, **16**, 1113–1117; (b)



- C. Giaginis and A. Tsantili-Kakoulidou, *J. Liq. Chromatogr. Relat. Technol.*, 2008, **31**, 79–96; (c) C. My Du, K. Valko, C. Bevan, D. Reynolds and M. H. Abraham, *Anal. Chem.*, 1998, **70**, 4228–4234.
- 21 (a) S. Bosisio, A. S. J. S. Mey and M. J. Michel, *J. Comput.-Aided Mol. Des.*, 2016, **30**, 1101–1114; (b) C. C. Bannan, K. H. Burley, M. Chiu, M. R. Shirts, M. K. Gilson and D. L. Mobley, *J. Comput.-Aided Mol. Des.*, 2016, **30**, 927–944; (c) J. P. M. Jämbeck, F. Mocci, A. P. Lyubartsev and A. Laaksonen, *J. Comput. Chem.*, 2013, **34**, 187–197.
- 22 G. Gerogiokas, G. Calabro, R. H. Henchman, M. W. Y. Southey, R. J. Law and J. Michel, *J. Chem. Theory Comput.*, 2013, **10**, 35–48.
- 23 (a) J. Michel, R. H. Henchman, G. Gerogiokas, M. W. Southey, M. P. Mazanetz and R. J. Law, *J. Chem. Theory Comput.*, 2014, **10**, 4055–4068; (b) G. Gerogiokas, M. W. Y. Southey, M. P. Mazanetz, A. Heifetz, M. Bodkin, R. J. Law and M. J. Michel, *Phys. Chem. Chem. Phys.*, 2015, **17**, 8416–8426; (c) G. Gerogiokas, M. W. Y. Southey, M. P. Mazanetz, A. Heifetz, M. Bodkin, R. J. Law, R. H. Henchman and M. J. Michel, *J. Phys. Chem. B*, 2016, **120**, 10442–10452.

